BASIC AND APPLIED SOCIAL PSYCHOLOGY, 37:202–213, 2015

ISSN: 0197-3533 print/1532-4834 online
DOI: 10.1080/01973533.2015.1049349

Routledge
Taylor & Francis Group

The Mediation Myth

Rex B. Kline
Concordia University

The mediation myth is the false belief that mediation is actually estimated in the typical mediation analysis. This myth is based on a trifecta of shortcomings: (1) the typical mediation study relies on an inadequate design; (2) the researcher uses a flawed analysis strategy; and (3) there is scant attention to assumptions that are required when estimating mediation. These problems stem from overgeneralizing the classical product method for estimating mediation and overreliance on statistical significance testing as a decision criterion in mediation analysis. The goals of this article are to (1) raise awareness of these difficulties among researchers and (2) provide a roadmap about design and analysis options for a more rigorous and scientifically valid approach to mediation analysis.

Mediation is an important concept in many disciplines, including psychology, education, and medicine, among others. It is relevant when the goal is to understand how changes are transmitted from a casual variable through one or more intervening variables, or mediators, which in turn leads to changes in an outcome (Little, 2013). Pearl (2014) described mediation analysis as telling us how nature works through the estimation of both direct and indirect effects of cause on outcome, where the latter involve the mediator(s). Relative magnitudes of direct versus indirect effects may have treatment or policy implications. For example, patient adherence is a key mediator between medical practice and patient outcomes (Fincham, 2007); thus, it is crucial to know for which particular treatments and disorders there are appreciable mediating effects of patient adherence. In such cases, efforts to boost patient compliance may improve outcomes even more than altering the treatment.

Mediation analysis dates to the theory of path coefficients by the geneticist Sewall Wright (1923, 1934), who demonstrated how to express coefficients for direct or indirect causal effects in terms of the data, or sample covariances among observed variables. Path analysis is nowadays considered as part of the family of techniques known as structural equation modeling (SEM), and direct versus indirect effects of observed or latent variables are routinely estimated in SEM (Kline, 2016). Other statistical frameworks for mediation analysis include inverse probability weighting (Robins, Hernán, & Brumback, 2000), the counterfactual or potential outcomes model (Rubin, 2005), principal stratification with propensity scores (Jo, Stuart, MacKinnon, & Vinokur, 2011), instrumental variable methods (Angrist, Imbens, & Rubin, 1996), conditional process modeling (Hayes, 2013), and causal mediation analysis (Pearl, 2014). Some of these methods are mentioned later when discussing contemporary options for mediation analysis.

Among works in which the product method, or the classical regression approach, for estimating indirect effects among continuous variables in linear parametric models where no interactions are assumed (Alwin & Hauser, 1975; Finney, 1972), Baron and Kenny (1986) is the best known and most widely cited example. By the year 2014, this particular article was cited more than 40,000 times (Kenny, 2014), and it was the single most cited article in the *Journal of Personality and Social Psychology* (MacKinnon & Pirlott, 2015). It is fair to say that Baron and Kenny and related works piqued an interest in mediation that continues to this day. Whether the enthusiasm for estimating mediation is backed up by proper research designs and analysis strategies is another matter as we shall see.

Correspondence should be sent to Rex B. Kline, Concordia University, 7141 Sherbrooke Street West, Montréal, QC H4B 1R6, Canada. E-mail: rex.kline@concordia.ca

The rationale of the product method is briefly explained for the basic mediation model presented in Figure 1a, where X, M, and Y designate, respectively, the cause, mediator, and outcome, and where a, b, and c' represent the path coefficients for the three direct effects in the model. Baron and Kenny (1986) described the four requirements for establishing mediation that are listed next and discussed afterward:

1. The cause affects the outcome ignoring the mediator, that is, coefficient c in Figure 1b is not zero.
2. The cause affects the mediator, that is, coefficient a in Figure 1a is not zero.
3. The mediator affects the outcome controlling for the cause, that is, coefficient b in Figure 1a is not zero.
4. To claim that the mediator is completely responsible for the association between cause and outcome, coefficient c' in Figure 1a should be zero.

In Figure 1a, the product ab estimates the indirect effect of X on Y through M. The quantity $ab + c'$ estimates the total effect of X on Y, or the sum of the direct and indirect effects of X. It also equals coefficient c in Figure 1b where X is the sole measured cause of Y. For continuous variables, $c - c' = ab$, that is, the difference between the total effect of X in Figure 1b ignoring M and the direct effect of X in Figure 1a controlling for M equals the product estimator of the indirect effect. The equality just mentioned does not hold for models where Y is a binary outcome and when using logistic regression; see Valeri and VanderWeele (2013) for more information.

The first requirement in the list just presented that coefficient c in Figure 1b should not be zero is problematic because it does not allow for inconsistent mediation where the signs of ab and c' in Figure 1a are different. In this case, the total effect c in Figure 1b could be zero or close to zero even though the magnitude of the product ab is appreciable. This could happen if the mediator acts as a suppressor variable such that the sign of coefficient b or c' in Figure 1a is the opposite of that of the corresponding Pearson correlation between Y and X or M; see MacKinnon, Fairchild, and Fritz (2007) for more information about inconsistent mediation. For these reasons, most researchers nowadays do not apply the first requirement, so next we consider only Requirements 2 to 4 from the original list.

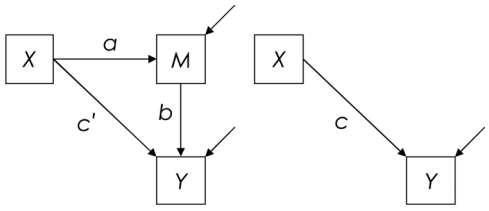

FIGURE 1 (a) Basic model of mediation for continuous variables with path coefficients where the indirect effect is estimated as the product ab. (b) Model with the same cause and outcome but no mediator.

Baron and Kenny (1986) stated requirements for establishing mediation in terms of zero versus nonzero coefficients, *not* in terms of statistical significance. One reason is that very small coefficients can be significant in large samples just as very large coefficients can be not significant in small samples. Another is that collinearity between X and M can reduce the power of tests for coefficients a, b, and c' in Figure 1a (Kenny, 2014). Nevertheless, most researchers use significance testing as a decision criterion in mediation analysis (Frazier, Tix, & Barron, 2004). For instance, some researchers look for both coefficients a and b in Figure 1a to be significant in order to claim that at least part of the effect of X on Y is transmitted through M. There are also ways to test the product estimator ab for significance, including an approximate test by Sobel (1982) that is amenable to hand calculation for indirect pathways with a single mediator and more contemporary methods based on bootstrapping (Preacher & Hayes, 2008). The observation that ab is significant would be taken by many researchers as evidence for mediation. But whether results of significance testing in mediation analysis generally have any meaningful interpretation is a critical question, one that is rarely considered in most mediation studies. We address this issue momentarily.

Mediation analysis depends on many assumptions that are rarely acknowledged in the typical empirical study. Some of these assumptions are also untestable in that the data can tell us nothing about whether the assumption is tenable or not tenable (Bullock, Green, & Ha, 2010). The overriding supposition is that of modularity. This means that the causal process is made up of components that are potentially isolatable, and thus can be analyzed as separate entities; that is, the causal process is not organic or holistic (Knight & Winship, 2013), and thus inseparable into parts. There is no empirical test of modularity; therefore, it must be assumed in mediation analysis, but few authors of mediation studies even mention this issue.

The hypothesis of mediation assumes that all corresponding directionality specifications are correct. In Figure 1a, for example, it is assumed that X is a cause of both M and Y and that Y is also caused by M. If the specifications just stated are incorrect, the quantity ab may have no meaningful interpretation. For example, if Y

really causes M in Figure 1a, or if this pair of variables mutually cause each other (feedback), then ab does not estimate the indirect effect of X on Y through M. Some researchers falsely believe that directionality specifications are actually tested in mediational analysis in particular and in SEM in general (Kline, 2016). The truth is that the model with its directionality specifications is assumed to be correct at the beginning of the analysis. Explanation of the fact just stated has to do with equivalent models, which are considered later, but directionality is generally assumed, not actually tested in mediation analysis.

The model in Figure 1a assumes that there are no unmeasured common causes (confounders) for any pair of variables among X, M, and Y. This includes the assumption of independent errors for variables M and Y. It is also assumed for Figure 1a that no omitted confounder of the association between M and Y is caused by X. The assumption of no interaction can be relaxed in some newer methods for estimating mediation. One is causal mediation analysis in which interaction between the cause and mediator is routinely estimated. In contrast, the classical product method assumes that any cause–mediator interaction is zero. This assumption is sometimes unrealistic. For example, the effect of patient compliance on outcome may depend on treatment; specifically, aversive side effect associated with a particular treatment may alter the role of patient compliance in recovery. Causal mediation analysis also supports evaluation of the assumption of no omitted causes of M and Y. Pearl (2014) elaborated on the assumptions for mediation analysis just described.

Finally, the model in Figure 1a assumes that scores on variable X have no measurement error, that is, $r_{XX} = 1.0$, where r_{XX} is a reliability coefficient. This is because exogenous variables in path models have no error terms; thus, there is no "room" in the model to accommodate $r_{XX} < 1.0$. Instead, score imprecision in X tends to show up "downstream" in the model, or in path coefficients for variables affected by X or in their error terms. If X were an experimentally manipulated variable, such as random assignment of cases to either a treatment or control condition, the assumption of no error variance may be plausible, but this is not generally true if X represents a measured (nonexperimental) attribute, especially an individual difference variable assessed through a self-report questionnaire. In contrast, because random error in measured endogenous variables is manifested in their disturbances, it is *not* assumed for M and Y in Figure 1a that $r_{MM} = r_{YY} = 1.0$. 1.0. There are ways in SEM to explicitly control for measurement error in each and every observed variable (Hayduk & Littvay, 2012), but such methods are infrequently used when mediational path models like Figure 1a are analyzed. Cole and Preacher (2014) described complications that can arise when appreciable measurement error is not controlled in path analysis. They included over- or underestimation of causal effects and reduction of statistical power that prevents rejection of false models among other problems that are generally magnified as models become more complex.

Assumptions in mediation analysis are addressed mainly through specification, design, and measurement. As said, the hypothesis that the cause and mediator interact can be evaluated with the data, but otherwise analysis by itself is rarely sufficient to establish causation. With a good knowledge of theory (specification), a strong research design that supports casual inference, and selection of measures with good psychometrics, fewer assumptions may be required, but as Pearl (2000) reminded us, "causal assumptions are prerequisite for validating any casual conclusion" (p. 136). This statement echoes that of Wright (1923) from nearly 80 years earlier when he said that "prior knowledge of the causal relations is assumed as prerequisite in the theory of path coefficients" (p. 240). Whether unverifiable assumptions are tenable is thus a matter of argument, not statistics, while researchers should use optimal research designs and measures.

PROPER RESEARCH DESIGNS FOR MEDIATION ANALYSIS

Most studies in the mediation literature are based on cross-sectional designs where all variables are concurrently measured. Such designs have no formal elements that directly support causal inference. The absence of time precedence in such designs is especially critical: If variables X, M, and Y in Figure 1a are measured at the same occasion, there is no way to establish which of two variables, a presumed cause and a presumed effect, occurred first. Therefore, the sole basis for causal inference in cross-sectional designs is assumption, one supported by a convincing, substantive rationale for specifying that, for instance, M causes Y in Figure 1a instead of the reverse or that M and Y mutually cause each other. This process relies heavily on the researcher to rule out alternative explanations of the association between M and Y. Doing so requires solid knowledge about the phenomenon under study. Sometimes alternative causal effects in cross-sectional designs can be ruled out by the nature of the variables. Suppose that X in Figure 1a is gender and that both M and Y are individual difference variables (e.g., patient adherence, medical status). It would be illogical to assume that M or Y could change a person's gender, so the alternative specifications $M \rightarrow X$ and $Y \rightarrow X$ would be indefensible.

It is otherwise usually possible in cross-sectional designs to generate equivalent versions of a path model

that explain the data just as well as the original model but make contradictory claims about directionality. Presented in the upper left part of Figure 2 is the basic mediation model we have considered to this point. The 17 other models in the figure are equivalent versions with identical fits to the same data. All models in the figure would have perfect fit because their degrees of freedom are zero, but the point is that analysis cannot ascertain which (if any) of these equivalent models is correct. For example, the models in the first row of Figure 2 are recursive and arbitrarily switch the roles of cause, mediator, and outcome among variables *X*, *M*, and *Y*. Models in the second row are nonrecursive and feature an equality-constrained direct feedback loop where the two variables in the loop each mediate the effect of the other. Instrumental variables must be added to any of the models in the second row in order to specify that each direct effect in the feedback loop is a free parameter or to add a disturbance covariance to the model (Rigdon, 1995). Models in the third row of Figure 2 with disturbance covariances are recursive and do not specify mediation, but they explain the data just as well as any hypothesis of mediation.

An appropriate design for estimating mediation has time precedence where the cause is measured before the mediator, which in turn is measured before the outcome. This requirement is consistent with Little's (2013) definition of mediation given at the beginning of this article that emphasizes the transmission of *changes* from cause to mediator to outcome. It also explains the distinction between the terms "indirect effect" and "mediation." Specifically, mediation always involves indirect effects, but not all indirect effects automatically signal mediation. This is especially true in cross-sectional designs with no time precedence, which do not allow or control for changes in any variable. *Thus, without a proper research design—one with time precedence—use of the term "mediation" is unwarranted.* The term "indirect effect" may still apply if directionality assumptions in a cross-sectional study have a strong rationale, but the term "mediation" should be reserved for research designs with formal elements that directly support causal inference.

A minimal design for mediation analysis that corresponds to Figure 1a is a measurement-of-mediation model (Bullock et al., 2010), where (a) the cause is an experimental variable where cases are randomly assigned to conditions, such as $X = 0$ for control and $X = 1$ for treatment; (b) mediator *M* is an individual difference variable that is not manipulated and is measured at a later time; and (c) and outcome *Y* is measured at a third occasion. Selecting the appropriate measurement schedule, or time lags, is critical. For example, measuring *M* too soon may not give the treatment enough time to affect *M*, but measuring *M* too late can miss temporary treatment effects that have dissipated. An advantage is that because *X* is an experimental variable, over replications it will be isolated from confounders that also affect *M* or *Y*. But because *M* is not manipulated, it may be plausible that *M* and *Y* share at least one unmeasured cause (i.e., their error terms are correlated); if so, then the coefficient for the path $M \rightarrow Y$ is actually not identified.

One way to identify the coefficient for the path $M \rightarrow Y$ in a measurement-of-mediation model is to specify an instrument for the mediator, such as a variable

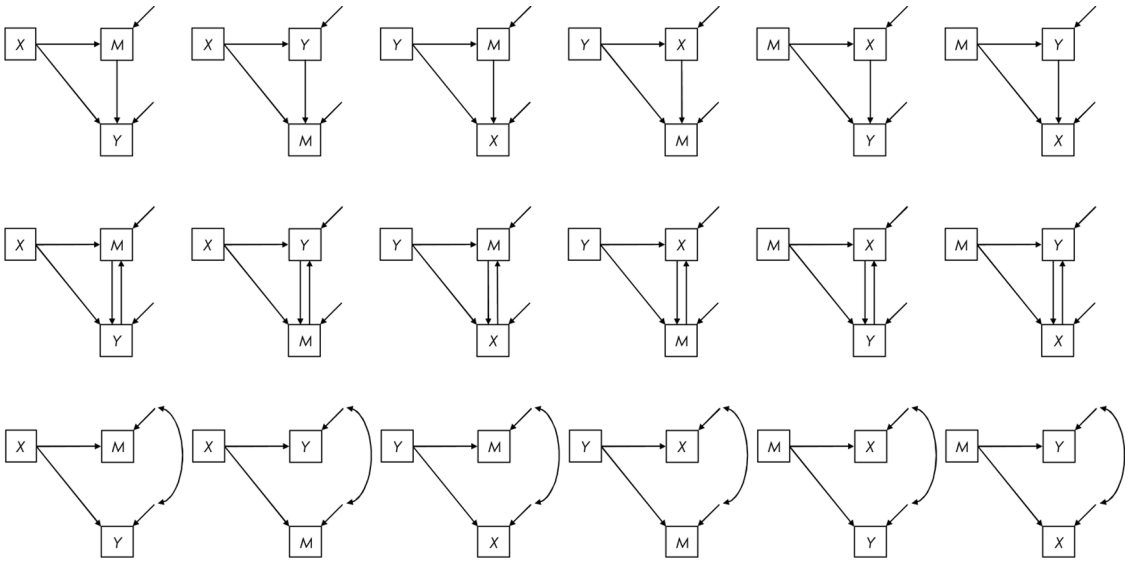

FIGURE 2 Equivalent models for three variables in cross-sectional designs without temporal precedence and without a strong rationale for directionality specification. *Note.* Reciprocal direct effects are constrained to equality.

that directly affects M but not Y and is also unrelated to the disturbance of Y (Antonakis, Bendahan, Jacquart, & Lalive, 2010). Exogenous variables make ideal instruments because by definition they are unrelated to all disturbances in the model. An example is a manipulated instrument where cases are randomly assigned to conditions that should directly affect M but not Y. Suppose that patients in a measurement-of-mediation study are randomly assigned to conditions that offer varying levels of incentive (including none) for adherence to treatment. This randomized manipulation to change M may indirectly affect Y, but over replications there should be no direct effect. Instruments can also be measured variables, but finding nonexperimental instruments can be challenging. In designs with a single mediator, it must be assumed that M completely mediates the relation between X and Y. This is a strong assumption that is not always convincing. See MacKinnon and Pirlott (2015) for additional discussion of instrumental variable methods in mediation analysis.

Imai, Keele, and Tingley (2010) described methods to evaluate the assumption of no unmeasured confounder of the relation between M and Y. It is a kind of sensitivity analysis where the researcher can explore and plot how the estimated mediation effect would change allowing for a confounder to theoretically change the extent of overlap in the error variances of M and Y. If the estimated change is appreciable due to a possible confounder, less confidence is warranted in the results. The method works by increasing the correlation between the error terms of M and Y and then evaluating the degree to which the estimated mediation effect changes. It can also generate a plot of how much the estimate changes, given a particular effect of the confounder. There are similar methods in regression analysis for evaluating the potential effect of unmeasured predictors that covary with measured predictors, or left-out variables error (Mauro, 1990). MacKinnon and Pirlott (2015) described other methods to evaluate the extent of confounding bias in mediation analysis.

A stronger experimental design is the manipulation-of-mediation model (Bullock et al., 2010), where the mediator is also a manipulated variable. Examples of mediators that are potentially manipulable include self-efficacy, goal difficulty, performance norms, and arousal, among others, but manipulating other kinds of internal states or situational factors may be difficult or unethical in some cases (Stone-Romero & Rosopa, 2011). It must be assumed in experimental mediational designs that manipulation affects just the mediator in question and not other mediators. This requires manipulations that target specific mediators, and meeting this requirement can be tricky given multiple mediators. The results may apply only to the subset of participants who responded to the manipulation of the mediator and not to the whole sample. Challenges of manipulating mediators are substantial, but successfully doing so isolates over replications confounders of X or M and Y. It also identifies the coefficient for the path $M \rightarrow Y$ without an explicit instrument.

Time precedence is also part of longitudinal designs for mediation described by Cole and Maxwell (2003), Maxwell and Cole (2007), Selig and Preacher (2009), Little (2013), and others. The smallest mediation design with repeated measures is the half-longitudinal design, which is depicted in Figure 3. The mediator and outcome are measured at both times, but the cause is measured only at Time 1 (see the figure). At Time 2, the disturbances of the mediator and outcome are assumed to covary. For continuous variables and assuming no interactions, the indirect effect is estimated as the product of coefficients a and b from the cross-lagged paths, respectively,

$$X_1 \rightarrow M_2 \text{ and } M_1 \rightarrow Y_2$$

The product estimator ab thus controls for autoregressive effects of the mediator and outcomes on *themselves* over time, which correspond to the paths

$$M_1 \rightarrow M_2 \text{ and } Y_1 \rightarrow Y_2$$

In the full-longitudinal design, X, M, and Y are each measured over at least three different occasions. This larger design actually consists of two replications of the half-longitudinal design. It also yields multiple estimators of the indirect effect, including the quantity associated with the single contiguous pathway through which X can affect Y through M and proxy estimators that are not (Little, 2013). If the model is correctly specified, though, values of all the estimators just mentioned should be comparable within the limits of sampling error.

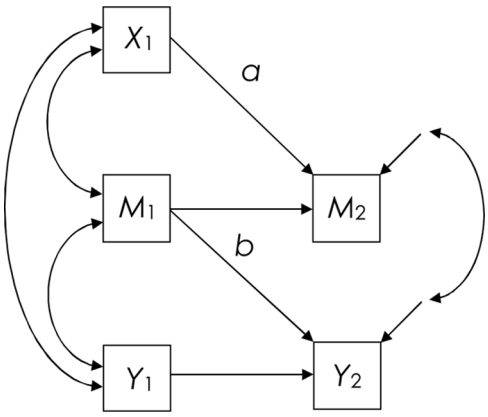

FIGURE 3 A half-longitudinal model of mediation where estimates of coefficients a and b control for autoregressive effects of, respectively, the mediator and outcome. *Note.* Subscripts indicate the time of measurement.

Measurement of variables at different times in longitudinal designs provides time precedence, and the hypothesis that *X* causes *Y* is bolstered if *X* were actually measured before *Y*. But time precedence is no guarantee. This is because the covariance between *X* and *Y* could still be relatively large even if *Y* causes *X* and the effect (*X*) is measured before the cause (*Y*). This can happen because *X* would have been affected by *Y* before either variable was actually measured in a longitudinal study. Longitudinal designs for estimating mediation assume optimal lags for the measurement occasions. Results of pilot investigations, such as distributed lag studies, may help to determine when an effect begins and when it ends (Little, 2013); otherwise, it can be difficult to specify the appropriate measurement schedule. Other challenges in longitudinal designs include case attrition and higher costs. This is probably why most mediational studies rely on concurrent rather than longitudinal measurement.

FOLLY OF STATISTICAL SIGNIFICANCE AS A DECISION CRITERION

The purpose of the analysis in too many studies is seen as discovering whether estimates of mediation are statistically significant or not statistically significant. In the eyes of most researchers, the former outcome is taken as positive evidence for mediation and gives a green light for elaborating on its meaning and implications. The latter outcome is not quite so joyous because it signals just the opposite. In this way, the dichotomous outcome of significance testing is used as a decision criterion in nearly all mediation analyses. For the reasons explained next, doing so is an exercise in folly not only in mediation studies but also in most other kinds of behavioral science research.

Significance testing is increasingly criticized in psychology, education, economics, and other disciplines as unscientific and unempirical (Kline, 2013; Lambdin, 2012). Other authors suggest that overreliance on significance testing leads to trained incapacity, or the inability of researchers to understand their own results or the status of a research literature due to (a) inherent limitations of significance tests and (b) and a myriad of associated cognitive distortions (Ziliak & McCloskey, 2008). Basic criticisms of significance testing are listed next and discussed afterward:

1. Outcomes of significance tests—*p* values—are wrong in most studies.
2. Researchers do not understand *p* values.
3. Significance tests do not tell researchers what they want to know.
4. Most applications of significance testing are incorrect.
5. Respecification in SEM based solely on significance testing merely capitalizes on chance.

Significance tests generally assume random sampling from known populations with no other source of error than sampling error. True random sampling is very rare, especially in human studies where most samples are ad hoc (convenience) samples made up of participants who happened to be available. Scores are affected by multiple types of error, including sampling error; measurement error; and, in treatment outcome studies, implementation error. Additional kinds include specification error, including left-out variables error, and a host of other threats to study internal, external, or conclusion validity. The idea that scores are typically affected by sampling error alone is a fantasy. Distributional assumptions of parametric significance tests, such as for normality or homoscedasticity, are implausible in many, if not most, studies. For instance, most empirical distributions are not even symmetrical, much less normal, and departures from both normality and homoscedasticity are often strikingly large (Keselman et al., 1998). Most researchers do not bother to verify distributional assumptions of significance tests because they falsely believe that such tests are robust even in small, unrepresentative samples (Hoekstra, Kiers, & Johnson, 2012).

Because *p* values are usually calculated under untenable assumptions, it is nearly impossible to believe that they are generally accurate. If this is true, then decisions based on *p* values are untrustworthy, too. Additional skepticism about significance testing is warranted in mediation analysis. For example, the Sobel (1982) test assumes normality, but product estimators of indirect effects do not generally follow normal distributions. Bootstrapped significance tests may be somewhat more accurate in large samples, but results from bootstrapping can be very biased in small samples, and most SEM studies are based on sample sizes that are too small (Westland, 2010). Bootstrapped significance tests still assume random sampling, which almost never happens in mediation studies—or in most other kinds of studies in the literature.

Another source of inaccuracy in *p* values is the type of null hypothesis tested, which is typically a nil hypothesis that some effect, difference, or association is zero in the population. The problem is that such hypotheses are known to be false in many research areas—including most mediation studies—before the data are even collected; that is, nil hypothesis are scientifically farfetched. Probabilities of data under implausible null hypotheses are too low, which make the data seem more

exceptional than they would be under a more plausible, non-nil hypothesis. The latter allows for an effect to be nonzero in the population. Non-nil hypotheses are more realistic in research areas where effect sizes are known not to be zero, but computer programs for statistical analysis almost always test nil hypotheses.

There is ample evidence that researchers do not understand p values. For example, about 80%–90% of psychology professors endorse multiple false beliefs about statistical significance, no better than psychology undergraduate students in introductory statistics courses (Haller & Krauss, 2002). Most errors involve overinterpretation that favor the researcher's hypotheses, which is a form of confirmation bias. For the case $\alpha = .01$ and $p < .01$, many researchers would conclude that the result is very unlikely to be due to sampling error, the likelihood that a Type I error was just committed is just as unlikely, and the likelihood that the null hypothesis is true is very low (<1% chance for all three). Many researchers would also conclude that the alternative hypothesis is very likely to be true and that the result is very likely to replicate (>99% chance for both). Unfortunately, *none* of these conclusions is correct, and there many other false beliefs about p values not mentioned here. Indeed, there are so many cognitive distortions in significance testing that one could argue that data analysis is based more on wishful thinking (what researchers would like to know) than on fact (what p values actually tell us).

The problem that many researchers do not generally verify distributional assumptions of significance tests was mentioned. If assumptions are checked, the wrong methods are typically used (Hoekstra et al., 2012), including significance tests that supposedly verify distributional assumptions of other significance tests, such as Levene's test for homoscedasticity. The problem with such tests is that their results are often wrong due in part to their own unrealistic assumptions (Erceg-Hurn & Mirosevich, 2008). Another issue is the widespread failure of researchers to report estimates of the a priori power of their significance tests. Without such estimates, it is impossible to interpret results that are not statistically significant. Specifically, if power is low, such results are expected, but if power is high, then the failure to obtain statistical significance has greater import. There are ways to estimate power in mediation analysis (Fritz & MacKinnon, 2007), but it is rare to find power estimates in reports of mediation studies.

Respecification of structural equation models based solely on statistical significance, such as adding paths with significant modification indexes or dropping paths with Wald W statistics that are not significant, does little more than capitalize on chance. Models so respecified may wind up fitting the data in a particular sample but are unlikely to replicate. In this way, the binary outcome of significance testing merely highlights a pattern of sample-specific results and tempts the researcher to believe otherwise (e.g., that the results are "real," or not due to sampling error). There is also the problem that basically *any* hypothesis can be supported (the result can be presented as statistically significant) through flexibility in data collection, analysis, or reporting (e.g., N, level of α) that is not always disclosed and is described by Simmons, Nelson, and Simonsohn (2011) as researcher degrees of freedom. The ease with which a false hypothesis can be supported in significance testing may fill the literature with false positive results. If such results are not replicated due to the erroneous belief that significant results are likely to replicate, then the original false claims will remain undetected.

Viewing the analysis as a search for statistical significance is a potential dead end not just in mediation studies but in most behavioral science research. In the former, researchers should estimate the magnitudes of mediated effects and evaluate whether such effects are large and precise enough to be appreciable in a particular context. There are ways to measure the relative sizes of direct versus indirect effects of both observed and latent variables in standardized metrics that are directly comparable across different studies (Lau & Cheung, 2012; Preacher & Kelley, 2011), when it is appropriate to analyze standardized effect size. There are also methods in mediation analysis for interval estimation, which allows the reporting of indirect effects with confidence intervals (Cheung, 2009). Because confidence intervals are calculated with the same standard errors that contribute to significance tests, researchers should be careful not to overinterpret the lower or upper bounds of confidence intervals (i.e., they are probably wrong). Specifically, confidence intervals should *not* be treated as significance tests in disguise. For example, the observation that a $(1 - \alpha) \times 100\%$ confidence interval includes zero has no special status in interval estimation because a result of zero is seen as just as plausible as any other value within the interval. The same result in significance testing would lead to the failure to reject the nil hypothesis that the corresponding parameter is zero at the alpha level of statistical significance for a two-tailed alternative hypothesis, but the value of zero may not appear in the corresponding confidence interval in a replication sample; that is, confidence intervals are subject to sampling error, too. The most important reform is replication: Mediated effects of appreciable magnitude and precision estimated in primary studies are not very interesting unless replicated. There is a paucity of meta-analytic summaries of mediation studies—see Shadish and Sweeney (1991)—but effect size estimation over replications is a standard part of meta-analysis.

The recent ban on significance testing in *Basic and Applied Social Psychology* (Trafimow & Marks, 2015)

is related to the issues just discussed. This policy is not as radical or as unprecedented as it may sound. For example, the reporting of p values was banned for a decade (1990–2000) in *Epidemiology* under the editorship of Kenneth Rothman, the journal's founder. The editor of *Memory & Cognition* from 1993–1997, Geoffrey Loftus actively discouraged the use of significance testing in guidelines for potential authors (Loftus, 1993). About two dozen research journals require the reporting and interpretation of effect sizes (Kline, 2013), which is consistent with the view the significance testing is insufficient to describe the results. It is more surprising that confidence intervals are also banned in *Basic and Applied Social Psychology*, but it is true that many of the same cognitive errors in significance testing are made about confidence intervals, too, a phenomenon called the law of diffusion of idiocy by Abelson (1997). Instead of a "crisis," editorial policies against significance testing present an opportunity for researchers to describe their results at a level closer to the data and to emphasize the *substantive significance* of their findings, not just their statistical significance. This includes mediation studies.

RELAXING THE ASSUMPTION OF NO INTERACTION

The classical product method for estimating indirect effects assumes that the cause X and mediator M do not interact. No interaction corresponds to effect homogeneity such that causal effects are constant across cases. In mediation analysis, this means that the effect of X on Y does not depend on the level of M just as the effect of M on Y has nothing to do with X. If there is cause–mediator interaction, the mediator is also a moderator that changes the effect of the X on Y, either attenuating or amplifying it, as a function of M. Because moderation is symmetrical, X also moderates the effect of M on Y when there is interaction between X and M.

Analyzing mediation and moderation in the same model estimates conditional casual effects, or effect heterogeneity. Two approaches to estimating mediation and moderation together are briefly described next, causal mediation analysis and conditional process modeling. Of the two, the former is better known in disciplines such as epidemiology and the latter is more familiar in the social sciences. Causal mediation analysis is the most general method. This is because direct, indirect, and total causal effects in causal mediation analysis are defined in a consistent way regardless of the statistical model, such as linear versus nonlinear, and regardless of possible interactions between causes and mediators. Causal mediation analysis can be extended to linear-, logistic-, log-linear, or Poisson-type regression analyses (e.g., odds ratios or risk ratios can be analyzed for dichotomous outcomes), and mediators can be continuous or dichotomous. Interaction effects between X and M are routinely estimated. If there is truly no cause–mediator interaction, causal mediation analysis and the classical product method for linear models with continuous outcomes yield the same estimates for the same data; otherwise, the two approaches can generate quite different results (Valeri & VanderWeele, 2013).

Causal mediation analysis is based on Pearl's (2014) mediation formula, which defines direct and indirect effects from a counterfactual perspective. For a randomized cause (e.g., treatment vs. control) and continuous mediator and outcome variables, the controlled direct effect is defined as the average difference between the treated and untreated cases if the mediator were controlled at the same level for all cases in the population. There is a different value of the controlled direct effect for each level of the mediator, so for a continuous M there are infinitely many controlled direct effects. The natural direct effect is the average difference in outcome if the causal variable were allowed to change from control to treatment, but the mediator is kept to the level that it would have taken in the control condition. Unlike the case for the controlled direct effect, the level of the mediator is not fixed to the same constant for all cases. Instead, the mediator is allowed to vary, but only over values that would be naturally observed in the control condition. If there is no interaction, then estimates of the controlled direct effect and the natural direct effect are equal for continuous variables in a linear model.

The parallel to the natural direct effect is the natural indirect effect. It estimates the amount of change among treated cases as the mediator changes from values that would be observed in the control group to the values it would obtain in the treatment group. In other words, the outcome is influenced by the cause due solely to its influence on the mediator. The total causal effect is the sum of the natural direct effect and the natural indirect effect. In contrast, the controlled direct effect does not a simple additive relation with either the natural direct effect or natural indirect effect, so it is not part of an effect decomposition in causal mediation analysis.

This example by Petersen, Sinisi, and van der Laan (2006) may help to clarify the counterfactual definitions just stated: Suppose that $X = 1$ is an antiretroviral therapy for HIV-infected patients and $X = 0$ is control. The mediator M is the blood level of HIV (viral load), and outcome Y is level of CD4 T-cells, or helper white blood cells. It is expected that the treatment effect depends on the patient's viral load. The controlled direct effect of treatment is the difference in average CD4 T-cell count if viral load were controlled at a single level for all cases. In contrast, the natural direct effect is the impact of treatment on CD4 T-cell count as it would have been observed if viral load were as in the control condition.

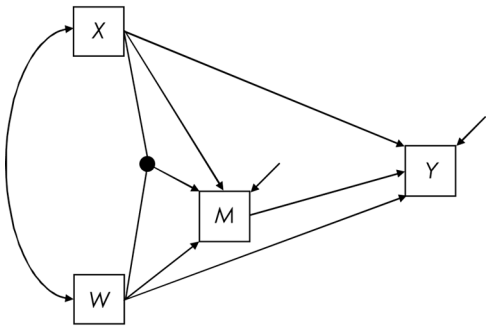

FIGURE 4 Model of mediated moderation and first-stage moderation where $X \rightarrow M$ depends on W but $M \rightarrow Y$ does not depend on W.

The natural indirect effect is the change in CD4 T-cell count in treatment if viral load shifted to what it would be without treatment to the level under treatment. The total effect of treatment on CD4 T-cell count is the sum of its natural direct and indirect effects; see the Appendix for a numerical example.

Cause–mediator interaction can also be estimated in conditional process modeling. Effects in this method are not defined in terms of counterfactuals, so it does not offer a common definition of mediation over linear versus nonlinear models. Two key concepts in conditional process modeling are those of mediated moderation and moderated mediation, also known as a conditional indirect effect. Both types of conditional causal effects just mentioned are represented in Figure 4 where the interaction between causal variables X and W is represented by the symbol for a closed circle. The interactive effect of X and W on Y is specified in Figure 4 as entirely indirect through M (mediated moderation). Due to the same interactive effect, the first path of the indirect effect of X on Y, or $X \rightarrow M$, depends on W, but the second path of the same indirect effect, or $M \rightarrow Y$, does not depend on W. Because moderation is symmetrical, it is also true in Figure 4 that the first stage of the indirect effect of W on Y, or $W \rightarrow Y$, depends on X, but the second stage, or $M \rightarrow Y$, does not depend on X. This pattern of moderated mediation is described by Edwards and Lambert (2007) as first-stage moderation. There are other patterns, such as second-stage moderation where just the second path of an indirect effect depends on an external variable, but all kinds of moderated mediation refer to conditional indirect effects. Preacher, Rucker, and Hayes (2007) described analysis strategies in conditional process modeling.

SUMMARY AND A ROADMAP TO BETTER MEDIATION ANALYSIS

There are hundreds and perhaps even thousands of published studies in which mediation is purported to be analyzed. Most of these studies are based on designs or analysis strategies that are inadequate to establish mediation, so relatively little of the extant literature on mediation is actually worthwhile. But mistakes of the past should help to guide better practice in the future. Accordingly, recommendations for better mediation analyses are given next; see the other articles in this special issue for additional recommendations.

Most mediation analyses are based on cross-sectional designs, which lack time precedence. This means that mediation actually cannot be established in such designs unless there is a convincing rationale for directionality specifications. Better designs include experimental mediational designs where the cause is a manipulated variable or longitudinal mediation designs where presumed causes, mediators, and outcomes are measured on different occasions. Experimental manipulation of the cause isolates that variable from confounders of the cause–outcome relation and from the cause–mediator relation, but there may be confounders of the mediator–outcome relation if the former is a measured, not manipulated variable. There are methods for conducting a sensitivity analysis of the assumption of no confounders of the mediator–outcome relation. If the result of such analyses is that the estimated mediation effect would change by an appreciable amount, given a confounder, then the interpretation of other findings as evidence for mediation may be unwarranted. Experimental designs are stronger when the mediator can also be manipulated or there is an instrument for a measured mediator. Both variations just mentioned can help to control for unmeasured causes of the mediator and outcome. Longitudinal mediation designs feature time precedence, but they guarantee neither the correctness of directionality specifications nor the absence of confounders. So there is no substitute for strong knowledge, good study planning, and reasonably correct specification regardless of the design.

Classical estimation methods assume no cause–mediator interaction or any other interactive effects that make a direct effect in a larger indirect pathway contingent on an external variable. There are ways in both casual mediation analysis and conditional process modeling to explicitly estimate cause–mediation interaction or conditional indirect effects where mediation depends on external variables. These methods may give more accurate results if the assumption of no interactions is untenable. Estimating interaction requires large samples and measures with very reliable scores, such as $r_{XX} > .90$; otherwise, the estimates can be very imprecise (Edwards, 2009). Even when the no interaction assumption is plausible, the failure to control measurement error in mediation analysis can seriously bias the results.

Statistical significance is too often taken as a "gold standard" for establishing mediation, but this decision

criterion is extreme flawed. This is because *p* values in mediation analyses may be especially inaccurate due to unrealistic distributional assumptions or sample sizes that are too small. Even if *p* values were generally correct, most researchers overinterpret them by seeing more meaning in *p* values than they actually convey. Thus, finding that an estimate of mediation is "significant" versus "not significant" means practically nothing by itself. The researcher should also demonstrate that the estimated mediation effect is large and precise enough to be meaningful in a particular context. Doing so requires thorough knowledge of a research area along with an appreciation of typical effect sizes. It is easier not to think about the problem by relying on automated decision rules (significance testing) that are blindly applied across all research areas. But mediation analysis that is scientifically relevant is neither simple nor easy. This is why the title of the recent article by Bullock et al. (2010) cautions the reader not to expect an easy answer to the question in mediation analysis, What is the causal mechanism? Easy answers clutter the literature on mediation. It is time to do better.

REFERENCES

Abelson, R. P. (1997). A retrospective on the significance test ban of 1999 (If there were no significance tests, they would be invented). In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 117–141). Mahwah, NJ: Erlbaum.

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, *40*, 37–47. doi:10.2307/2094445

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455. doi:10.1080/01621459.1996.10476902

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*, 1086–1120. doi:10.1016/j.leaqua.2010.10.010

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. doi:10.1037/0022-3514.51.6.1173

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*, 550–558. doi:10.1037/a0018933

Cheung, M. W. L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, *41*, 425–438. doi:10.3758/BRM.41.2.425

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*, 558–577. doi:10.1037/0021-843X.112.4.558

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*, 300–315. doi:10.1037/a0033805

Edwards, J. R. (2009). Seven deadly myths of testing moderation in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 143–164). New York, NY: Taylor & Francis.

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. doi:10.1037/1082-989X.12.1.1

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63*, 591–601. doi:10.1037/0003-066X.63.7.591

Fincham, J. E. (2007). *Patient compliance with medications: Issues and opportunities*. Binghamton, NY: Pharmaceutical Products Press.

Finney, J. M. (1972). Indirect effects in path analysis. *Sociological Methods & Research*, *1*, 175–186. doi:10.1177/004912417200100202

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, *51*, 115–134. doi:10.1037/0022-0167.51.1.115

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*, 233–239. doi:10.1111/j.1467-9280.2007.01882.x

Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online*, *7*, 1–17. Retrieved from http://www.dgps.de/fachgruppen/methoden/mpr-online/

Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*, *12*(159). doi:10.1186/1471-2288-12-159

Hayes, A. F. (2013). *Introduction to mediation, moderation, and process control analysis: A regression-based approach*. New York, NY: Guilford Press.

Hoekstra, R., Kiers, H. A. L, & Johnson, A. (2012). Are assumptions of well-known statistical techniques checked, and why (not)? *Frontiers in Psychology*, *3*, Article 137. Retrieved from http://journal.frontiersin.org/article/10.3389/fpsyg.2012.00137/full

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*, 309–334. doi:10.1037/a0020761

Jo, B., Stuart, E. A., MacKinnon, D. P., & Vinokur, A. D. (2011). The use of propensity scores in mediation analysis. *Multivariate Behavioral Research*, *46*, 425–452. doi:10.1080/00273171.2011.576624

Kenny, D. A. (2014). *Mediation*. Retrieved from http://davidakenny.net/cm/mediate.htm#CI

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., … Levin, J. R. (1998). Statistical practices of education researchers: An analysis of the ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, *68*, 350–368. doi:10.3102/00346543068003350

Kline, R. B. (2013). *Beyond significance testing: Statistics reform in the behavioral sciences* (2nd ed.). Washington, DC: American Psychological Association.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.

Knight, C. R., & Winship, C. (2013). The causal implications of mechanistic thinking: Identification using directed acyclic graphs (DAGs). In S. L. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 275–299). New York, NY: Springer.

Lambdin, C. (2012). Significance tests as sorcery: Science is empirical—significance tests are not. *Theory & Psychology*, *22*, 67–90. doi:10.1177/0959354311429854

Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods*, *15*, 3–16. doi:10.1177/1094428110391673
Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.
Loftus, G. R. (1993). Editorial comment. *Memory & Cognition*, *21*, 1–3. doi:10.3758/BF03211158
MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. doi:10.1146/annurev.psych.58.110405.085542
MacKinnon, D. P., & Pirlott, A. G. (2015). Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. *Personality and Social Psychology Review*, *19*, 30–43. doi:10.1177/1088868314542878
Mauro, R. (1990). Understand L.O.V.E. (left out variables error): A method for estimating the effects of omitted variables. *Psychological Bulletin*, *108*, 314–329. doi:10.1037/0033-2909.108.2.314
Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*, 23–44. doi:10.1037/1082-989X.12.1.23
Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York, NY: Cambridge University Press.
Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods*, *19*, 459–481. doi:10.1037/a0036434
Petersen, M. L., Sinisi, S. E., & van der Laan, M. J. (2006). Estimation of direct causal effects. *Epidemiology*, *17*, 276–284. doi:10.1097/01.ede.0000208475.99429.2d
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/BRM.40.3.879
Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93–115. doi:10.1037/a0022658
Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227. doi:10.1080/00273170701341316
Rigdon, E. E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research*, *30*, 359–383. doi:10.1207/s15327906mbr3003_4
Robins, J. M., Hernán, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*, 550–560. doi:10.1097/00001648-200009000-00011
Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*, 322–331. doi:10.1198/016214504000001880
Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*, 144–164. doi:10.1080/15427600902911247
Shadish, W. R., & Sweeney, R. B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *Journal of Consulting and Clinical Psychology*, *59*, 883–893. doi:10.1037/0022-006X.59.6.883
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. doi:10.1177/0956797611417632
Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290–312). San Francisco, CA: Jossey-Bass.
Stone-Romero, E. F., & Rosopa, P. J. (2011). Experimental tests of mediation models: Prospects, problems, and some solutions. *Organizational Research Methods*, *14*, 631–646. doi:10.1177/1094428110372673
Trafimow, D., & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, *37*, 1–2. doi:10.1080/01973533.2015.1012991
Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*, 137–150. doi:10.1037/a0031034
Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications*, *9*, 476–487. doi:10.1016/j.elerap.2010.07.003
Wright, S. (1923). The theory of path coefficients: A reply to Niles's criticism. *Genetics*, *8*, 239–255.
Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*, 161–215. doi:10.1214/aoms/1177732676
Ziliak, S., & McCloskey, D. N. (2008). *The cult of statistical significance: How the standard error costs us jobs, justice, and lives*. Ann Arbor: University of Michigan Press.

APPENDIX: NUMERICAL EXAMPLE OF NATURAL DIRECT, INDIRECT, AND TOTAL EFFECTS

Presented next is a numerical example based on one by Peterson et al. (2006), where $X = 1$ is an antiretroviral therapy for HIV and $X = 0$ is control; the mediator M is the blood level of HIV (viral load); and the outcome Y is level of CD4 T-cells (helper white blood cells). The two unstandardized regression equations for this example are listed next:

$$\widehat{M} = \beta_0 + \beta_1 X$$
$$\widehat{Y} = \theta_0 + \theta_1 X + \theta_2 M + \theta_3 XM \qquad (1)$$

where $\beta 0$ and $\theta 0$ are the intercepts for, respectively, the regressions of M on X and of Y on X, M and the product term XM; the coefficient for X when predicting M is $\beta 1$; and θ_1–θ_3 are the coefficients for, respectively, X, M, and XM when predicting Y.

The controlled direct effect (CDE) of X is how much outcome Y would change on average if the mediator were controlled at the same level $M = \mathrm{m}$ for all cases, but the treatment were changed from $X = 0$ (control) to $X = 1$ (treatment). The natural direct effect (NDE) is how much the outcome would change on average if X were changed from control to treatment, but the mediator is kept to the level that it would have taken in the control condition. The natural indirect effect (NIE) is the amount the outcome would change on average in the treatment condition, but the mediator changes from as it would from the control condition to the treatment condition. The total effect of X on Y is the sum of NDE and NIE.

Given the expressions in Equation 1, the CDE, NDE, and NIE can be expressed as follows

(Valeri & VanderWeele, 2013):

$$\begin{aligned} \text{CDE} &= \theta_1 + \theta_3 m \\ \text{NDE} &= \theta_1 + \theta_3 \beta_0 \\ \text{NIE} &= (\theta_2 + \theta_3)\beta_1 \end{aligned} \quad (2)$$

In the equation just listed, note that the CDE is defined for a particular level of the mediator ($M = m$), and the NDE is defined at the predicted level of the mediator in the control group ($X = 0$). This predicted level is β_0, which is the intercept in the equation for regressing M on X. Also note that if there is no interaction, then $\theta_3 = 0$ (see Equation 1). In this case, both the CDE and NDE are equal to the direct effect assuming no interaction, or θ_1, and the NIE equals the classical product estimator of the indirect effect, or $\beta_1\theta_2$.

Suppose in a particular sample that the two unstandardized regression equations are

$$\begin{aligned} \widehat{M} &= 1.70 - .20X \\ \widehat{Y} &= 450.00 + 50.00X - 20.00M - 10.00XM \end{aligned} \quad (3)$$

In words, the predicted viral load in the control group is 1.70, but treatment reduces this count by .20. For control patients with no viral load, the predicted level of CD4 T-cells is 450.00. Treatment increases this count by 50.00 for patients with no viral load, and for every one-point increase in viral load for control patients, the level of CD4 T-cells decreases by 20.00. For treated patients, the slope of the regression line for predicting the level of CD4 T-cells from viral load decreases by 10.00 compared with control patients. These equations imply that

$$\begin{aligned} &\beta_0 = 1.70 \text{ and } \beta_1 = -.20 \\ &\theta_0 = 450.00, \theta_1 = 50.00, \theta_2 = -20.00, \text{ and } \theta_3 = -10.00 \end{aligned}$$

Continuing with the same example, the direct effect of treatment versus control at a given level of viral load $M = m$ is

$$\text{CDE} = 50.00 - 10.00\,m$$

The researcher can select a particular value of m and then estimate the CDE by substituting this value in formula just listed. Another option is to estimate the direct effect at the weighted average value of M for the whole sample. The direct effect of treatment estimated at the level of viral load that would have been observed in the control condition is

$$\text{NDE} = 50.00 - 10.00\ (1.70) = 33.00$$

where 1.70 is the predicted value of the viral load in the control condition ($X = 0$) (see Equation 17.9). The indirect effect of treatment allowing viral load to change as it would from the control condition to the treatment condition is estimated as

$$\text{NIE} = (-20.00 - 10.00) - .20 = 6.00$$

where $-.20$ is the difference in viral load between the control and treatment conditions (see Equation 17.9). The total effect (TE) of treatment is the sum of its natural direct and indirect effects, or

$$\text{TE} = 33.00 + 6.00 = 39.00$$

Thus, antiretroviral therapy increases the level of CD4 T-cells by 39.00 through both its natural direct effect (33.00) and its natural indirect effect through viral load (6.00).