General Article

Advances in Methods and Practices in Psychological Science
July-September 2025, Vol. 8, No. 3, pp. 1–19

DOI: 10.1177/25152459251355585
www.psychologicalscience.org/AMPPS

On Partial Versus Full Mediation and the Importance of Effect Sizes

Thomas Ledermann[1], Myriam Rudaz[1], and Matthew S. Fritz[2]
[1]Department of Human Development and Family Science, Florida State University, Tallahassee, Florida; and [2]Department of Educational Psychology, University of Nebraska–Lincoln, Nebraska

Abstract
Theoretical models involving one or multiple intervening variables often posit whether a cause influences an outcome both directly and indirectly or only indirectly. In testing mediation, this distinction of partial and full mediation has become a subject of debate because of statistical issues. We extend the critique on this notion and provide insights into what a statistically significant direct effect between a cause and an outcome in a mediation model can mean. We also evaluate different effect size measures for direct and indirect effects and offer practical recommendations for assessing mediation mechanisms, which we illustrate using different examples. The broader relevance of these recommendations beyond mediation analysis is discussed.

Keywords
theoretical mediation, statistical mediation, effect size, sensitivity analysis, power analysis, open data, open materials

Received 5/4/24; Revision accepted 5/25/25

Mediation analysis has a long history (e.g., Hyman, 1955; MacCorquodale & Meehl, 1948; Wright, 1921) and currently enjoys a high popularity in the behavioral and social sciences that can be traced back to works published in the 1980s by, among others, Judd and Kenny (1981a, 1981b), James and Brett (1984), and of course, Baron and Kenny (1986). Over the last decades, statistical techniques have been developed that facilitate the assessment of mediation in simple and complex models involving multiple mediators or predictors and continuous and categorical variables (e.g., Hayes, 2018; Iacobucci, 2008; MacKinnon, 2008; Muthén et al., 2016; Pearl, 2009; VanderWeele, 2015). These advances provide researchers with insights into mediation mechanisms in models of causal relations as no other method does.

A distinction often made in mediation models is the one between partial and full mediation, also called incomplete and complete mediation (e.g., James & Brett, 1984; Kenny et al., 1998; Mathieu & Taylor, 2006; Shrout & Bolger, 2002). Over the last decade, there has been a debate surrounding the virtues of this distinction; some have advocated that it should be abandoned completely because of issues with how this distinction is assessed statistically (e.g., Hayes & Preacher, 2014; Preacher & Kelley, 2011; Rucker et al., 2011). Yet this distinction continues to be made in methodological articles (e.g., Sim et al., 2022), theoretical models (e.g., Schmader & Sedikides, 2018), individual studies (e.g., Le et al., 2024), and meta-analyses (e.g., Tran et al., 2022). A PubMed search revealed that in 2024, the terms "partial mediation" and "full mediation" appeared in 111 and 44 publications, respectively.

In this article, we extend the discussion of the distinction between partial and full mediation and the importance of additional statistical analyses, especially the calculation of effect sizes. We begin by showing that theoretical models involving one or multiple mediators often imply whether a cause (antecedent) influences an outcome (consequent) both directly and indirectly or

Corresponding Author:
Thomas Ledermann, Department of Human Development and Family Science, Florida State University, Tallahassee, Florida
Email: tledermann@fsu.edu

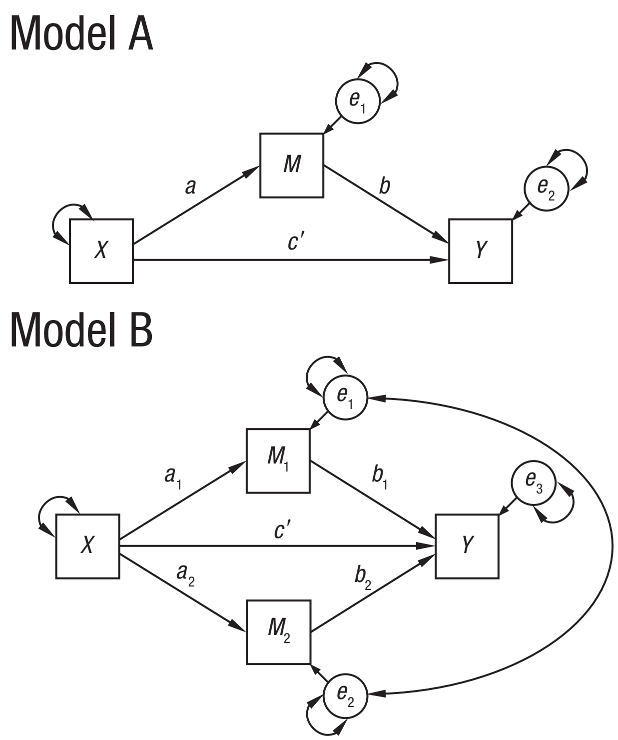

Fig. 1. Path diagrams of a simple mediation model (Model A) and a model with two simultaneous mediators (Model B). Rectangles represent manifest variables, circles represent residuals, single-headed arrows represent regression weights, double-headed arrows pointing to single variables represent variance parameters, and double-headed arrows pointing to different variables represent covariance parameters.

only indirectly. Next, we extend the critique of this distinction and explain why it is problematic statistically. We also elucidate what a significant direct effect between a cause and an outcome can mean and discuss requirements for causal relationships. We then evaluate effect size measures for both direct and indirect effects and provide equations to convert the direct effects. Finally, we offer practical recommendations for assessing mediation mechanisms, which we illustrate using hypothetical and real data.

Theoretical Perspective

The distinction between partial and full mediation can be found in many theoretical models that implicitly or explicitly posit that one variable causes another variable both directly and indirectly or only indirectly. Full mediation is indicated when a theoretical model states that a cause influences an outcome only indirectly through one or more mediators (James & Brett, 1984). An example is Schmader and Sedikides's (2018) conceptual model of state authenticity as fit to the environment. This framework posits that the fit between a person and the environment influences the person's motivation to approach situations indirectly through state authenticity, which the authors defined as a sense of being in alignment with one's own true self. Partial mediation is indicated when a theoretical model states that an antecedent influences an outcome both directly and indirectly through one or more mediators (James & Brett, 1984; Judd & Kenny, 1981b). For example, Karremans et al. (2017) hypothesized that mindfulness influences relational processes both directly and indirectly through awareness, emotion regulation, executive control, and self–other connectedness. Although often considered alone, theoretical models regularly combine partial and full mediation as smaller pieces of a larger causal model. A popular example is the theory of planned behavior (TPB; Ajzen, 1991). According to the TPB, people's attitude toward a planned behavior and the subjective norm predicts their behavior indirectly through their intention, whereas people's perceived behavioral control predicts their behavior both directly and indirectly through their intention.

These examples underscore that theoretical conceptualizations of causal frameworks often distinguish between partial and full mediation and that this distinction provides a more nuanced understanding of the process through which a cause is expected to affect an outcome. Once the mediation mechanism is identified, the distinction between partial and full mediation also provides researchers with insights into where it is appropriate to intervene to change the outcome (e.g., Ledermann & Macho, 2015; Loh et al., 2022). Although the statistical testing of partial versus full mediation is not without issues, this distinction provides easy to understand language to describe and discuss the mechanism by which a cause influences an outcome through one or more intervening variables.

Statistical Perspective

Consider the most basic mediation model that consists of a single cause, X, a single mediator, M, and a single outcome, Y. If M and Y are both continuous variables, then this simple mediation model (see Model A in Fig. 1) can be expressed by three linear equations:

$$M = i_1 + aX + e_1, \tag{1}$$

$$Y = i_2 + bM + c'X + e_2, \tag{2}$$

$$Y = i_3 + cX + e_3, \tag{3}$$

where a, b, and c' are estimates of the direct effects; i_1, i_2, and i_3 are intercepts; and e_1, e_2, and e_3 are residuals. The product ab is the estimate of the indirect or mediating effect of X on Y through M, and the sum of ab and c' is equal to the estimate of the total effect, c, such that $c = ab + c'$, assuming the relationships between the variables are linear and there are no missing data (MacKinnon et al., 1995). In this model, mediation is concluded to occur when ab is statistically significant and according to some authors (e.g., Fritz et al., 2012; Ledermann & Macho, 2009; Yzerbyt et al., 2018), when the two direct effects that make up this indirect effect are statistically significant (joint significance test; MacKinnon et al., 2002). The indirect effect and the direct effect c' are said to be consistent when ab and c' are both statistically significant and have the same sign (e.g., Kenny et al., 1998) and inconsistent when ab and c' have opposite signs (MacKinnon et al., 2000).

The simple mediation model can be easily expanded by adding more intervening variables. Model B of Figure 1 shows a model with two simultaneous or parallel mediators. In this model, there are two specific indirect effects, a_1b_1 and a_2b_2, a total indirect effect equal to the sum of the specific indirect effects, $a_1b_1 + a_2b_2$, a direct effect of X on Y (partialling out M_1 and M_2), c', and a total effect, c, equal to $a_1b_1 + a_2b_2 + c'$. Mediation is concluded to occur if one or both of the specific indirect effects are statistically significant or if the total indirect effect is statistically significant.

Now consider the case in which both a_1b_1 and a_2b_2 are statistically significant and have the same sign. Here, each mediator alone partially mediates the X to Y relation. Next, consider the case in which a_1b_1 and a_2b_2 are statistically significant but opposite in sign. This may sound like a contrived situation, but MacKinnon et al. (2001) found that the Athletes Training and Learning to Avoid Steroids (ATLAS) program (Goldberg et al., 1996) increased both reasons to avoid steroids and reasons to use steroids, which had opposite effects on intentions to use steroids. This resulted in specific indirect effects of opposite signs; the one through reasons to use steroids was positive, and the one through reasons to avoid steroids was negative. In a model with two simultaneous mediators, if the two specific indirect effects are equal in size but opposite in sign, they cancel each other out, resulting in a total indirect effect that is zero, and if c' is zero, then the total effect is zero as well. This example illustrates that mediation can exist in situations in which there is no total effect (see also MacKinnon et al., 2000; Shrout & Bolger, 2002). When c' is statistically significant and of the same sign as a_1b_1, then the effect through M_1 is consistent with c', and the effect through M_2 is inconsistent. If the two indirect effects cancel each other out because the magnitude of the two is equal, the total indirect adds up to zero, and the total effect becomes c'.

Issues in testing partial versus full mediation

In testing mediation, partial statistical mediation is said to occur when both the indirect effect and c' are significantly different from zero, and full statistical mediation is said to be indicated when the indirect effect is significantly different from zero but c' is not (e.g., Little et al., 2007; Mathieu & Taylor, 2006). This testing for partial versus full statistical mediation presents several issues. One is the reliance on the outcome of statistical null hypothesis significance testing (Hayes, 2018; Montoya & Hayes, 2017; Preacher & Kelley, 2011; Rucker et al., 2011; Wood et al., 2008). Assuming mediation is determined to be present, with at least one mediator showing a statistically significant specific indirect effect, then the distinction between partial and full mediation relies solely on the statistical significance of c'. Because significance tests are sample-size dependent, full statistical mediation could simply be due to a study being underpowered, which could be a result of measurement error (Fritz et al., 2016), to find a significant c' effect no matter how large this effect is, whereas partial mediation could be due to a large sample size so that c' is statistically significant no matter how small that effect is (e.g., Hayes, 2018; Shrout & Bolger, 2002). That is, full statistical mediation could be due to a Type II error, whereas partial statistical mediation could be due to a Type I error. Relying on significance tests for testing partial versus full mediation is further problematic because significance tests of indirect effects have been found to have generally more power than tests of the direct effect c' (Fritz & MacKinnon, 2007; Kenny & Judd, 2014) and because full statistical mediation is more likely to occur when the total effect is small (Preacher & Kelley, 2011).

A second issue concerns the practical implications of the distinction between partial and full mediation (Preacher & Kelley, 2011; Rucker et al., 2011). A researcher may be inclined to infer that results indicating full mediation suggest that a mediator is important and that results indicating partial mediation suggest that a mediator is less important. Such inferences are problematic, especially when relying solely on null hypothesis significance testing (Preacher & Kelley, 2011).

A third issue concerns the claim of full mediation. Claiming full mediation is tantamount to saying there is no direct effect between X and Y, which is essentially a claim of a null result. Claiming there is no effect can be problematic because there are almost always plausible alternative explanations that are often difficult to rule out because of the limitations of empirical studies (Wulff et al., 2023). These limitations include a lack of power to detect substantial effects, unreliable or invalid measures, undetected nonlinear relationships between the variables, violation of distribution assumptions, sampling error, the use of inappropriate statistical methods, or model misspecification. There may also be unmeasured subpopulations for which the null result is not true (see also Jacob et al., 2019; Landis et al., 2014). For example, the sample may consist of two subgroups, one group in which c' is negative and the other group in which c' is positive; together, the results may imply full statistical mediation, an illustration of Simpson's paradox (Shrout & Bolger, 2002). Thus, even if c' is zero, it does not equate to knowing that there is full mediation, since the absence of an effect cannot be asserted without ruling out all potential alternative explanations. Consequently, this issue renders any claim of full statistical mediation that is based on a single study problematic.

What does a statistically significant direct effect mean?

Finding a significant c' effect can be due to several factors. It may simply mean that the model misses one or more mediators, resulting in a misspecified model. For example, the stress-divorce model of Randall and Bodenmann (2009) proposes that the effect between everyday stress and mutual alienation is simultaneously mediated through time spent together, marital communication, physical and psychological problems, and problematic personality traits. The omission of one of these four hypothesized mediators from the analysis is likely to result in a c' effect that is statistically significant.

A significant c' effect may also be found because a confounder of the relationship between X and Y or between M and Y is omitted from the statistical analysis, which typically results in an overestimation of the indirect effect and direct effect c' (see also Loh et al., 2022). Omitting a mediator or a confounder of one of the direct effects violates the no-omitted-variable assumption (e.g., Fritz et al., 2016; Tofighi & Kelley, 2016; VanderWeele, 2010). Omitting a cause, either a mediator or a confounder of Y, tends to bias the results in favor of finding a significant c' effect (Bullock et al., 2010; Fritz et al., 2016; Shrout & Bolger, 2002).

It is important to note that it is never possible to know with certainty whether all relevant variables have been included in a model and that a c' effect that is negligible in size or even zero does not mean that all the relevant mediators and covariates have been included because there can be unmodeled competing mechanisms through which a cause influences an outcome, resulting in a true zero value of c'. Competing mechanisms can also occur in a model with two or more parallel mediators when one indirect effect is positive and the other one is negative, as in the MacKinnon et al. (2001) ATLAS example. In models with a single mediator and a single predictor, it is plausible that there is a direct effect between the predictor variable and the outcome from both a statistical and theoretical perspective because it is very likely that such a simple model lacks important variables, mediators and predictor variables.

Finally, a significant c' effect may simply be due to measurement error in the variables. In particular, measurement error in the mediator tends to lead to an underestimation of the b-path, which tends to attenuate ab and inflate c' (Baron & Kenny, 1986; Cole & Preacher, 2014; Fritz et al., 2016; Gonzalez & MacKinnon, 2021; Hoyle & Kenny, 1999, Ledgerwood & Shrout, 2011). Clearly, it cannot be emphasized enough that the use of reliable and valid measures is crucial in testing mediation mechanisms.

Causal Relationships

Theoretical models often posit causal relationships between constructs. Although randomized experiments are the "gold standard" for making causal inferences (cf. Berk, 2005), researchers often employ regression analysis and related techniques, such as multilevel modeling (MLM) and structural equation modeling (SEM), to test hypothesized causal relationships. Inferring a causal relationship between two variables requires the fulfillment of three widely accepted criteria (e.g., Kenny, 1979): There is an (observed) association between the variables (e.g., a substantial correlation), there is temporal precedence of the variables (X precedes Y in time), and the association is nonspurious (i.e., the association is not due to a confounding variable). Nonspuriousness is arguably the most challenging criterion (Rohrer et al., 2022), requiring a strong theoretical rationale for the hypothesized model and a correctly specified statistical model that includes all relevant covariates. Note that these requirements are necessary but not sufficient for inferring causation (see also Sobel, 1996). In ruling out alternative explanations, it is crucial to use reliable and valid measures, ensure that the sample size is adequate, employ statistical estimation methods that are appropriate for the data, and control for all potential confounders (see also MacKinnon, 2008).

In specifying the model to be estimated, Pearl (2001) delineated four critical assumptions important for inferring causal indirect effects: (a) no unmeasured confounders of the X-Y relationship, (b) no unmeasured confounders of the X-M relationship, (c) no unmeasured confounders of the M-Y relationship, and (d) no confounders directly affected by X that also affect the M-Y relationship. A violation of any of these assumptions is likely to lead to biased estimates of the indirect effect of interest and jeopardize causal inferences about the mediation mechanism.

For randomized designs, MacKinnon et al. (2020) discussed the testing of the effect of the XM interaction (i.e., the interaction between X and M) on Y. If the effect of this interaction on Y is zero, then the b and c' effects do not differ across conditions, and the indirect effect ab represents the causal indirect effect. If the effect of this interaction is nonzero, then b differs across levels of X, and c' differs across levels of M.

Table 1. Characteristics of Effect Sizes for Direct and Indirect Effects

	Direct effects			Indirect effects	
Characteristic	b_s	r_{sp}^2	Cohen's f^2	ab_s	υ
Easy to interpret	Yes	Yes	No	Yes	No
Benchmarks for small, medium, and large	No	No	Yes	No	No
Theoretical range	$-\infty$ to $+\infty$	0 to 1	0 to $+\infty$	$-\infty$ to $+\infty$	–1 to 1
Information about the direction	Yes	No	No	Yes	No
Convertible to r_{sp}^2 or Cohen's f^2	Yes	Yes	Yes	No	No
Applicable to complex mediation models	Yes	Yes	Yes	Yes	No
Useful for interaction and nonlinear effects	No	Yes	Yes	No	No
Parameters required	$b_s(b, SD_X, SD_Y)$	$r_{sp}^2(b_s, R_j^2)$	$f^2(r_{sp}^2, R_Y^2)$	$ab_s(ab, SD_X, SD_Y)$	$\upsilon(b_{b_s}^2, R_y^2, r_{XY}^2)$

Note: b_s = standardized estimate; b_{b_s} = standardized estimate of the *b*-path; r_{sp}^2 = semipartial correlation squared; ab_s = standardized effect estimate; R_Y^2 = proportion of explained variance in the outcome; R_j^2 = proportion of explained variance in the *j*th predictor variable by the other predictor variables; r_{XY} = correlation between *X* and *Y* in the simple mediation model.

Another recommendation is the use of sensitivity analysis to assess the robustness of the effects to potential omitted confounders or mediators (Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010). This type of analysis can be used to determine how large the effect of an unmeasured confounder would need to be on the mediator and the outcome to explain an estimated effect away. Current sensitivity-analysis techniques are best suited for assessing potential confounding effects that affect path-*b*. If there is an unmeasured confounder affecting both the mediator, *M*, and the outcome, *Y*, then the residuals of these two variables will be correlated. The size of this correlation is an indicator of how robust *b* is to omitted confounders affecting *M* and *Y*.

Recommendations

As illustrated, distinguishing between partial and full mediation in a theoretical model can have substantial value even though its testing presents challenges. One of the biggest issues is the reliance on statistical significance tests for c' (e.g., Hayes, 2018; Montoya & Hayes, 2017; Preacher & Kelley, 2011; Rucker et al., 2011; Wood et al., 2008). To address this issue, we recommend the practice of reporting and interpreting effect sizes and being cautious in claiming full mediation. We also discuss the consideration of sensitivity analysis, power analysis, and the testing of possible interaction effects when using randomized designs.

Reporting effect sizes

A first recommendation is to encourage the practice of reporting and interpreting effect sizes (e.g., American Psychological Association, 2020; Cumming, 2014). Several effect size measures have been proposed for assessing mediation (e.g., Lachowicz et al., 2018; MacKinnon, 2008; MacKinnon et al., 2007; Preacher & Hayes, 2008; Preacher & Kelley, 2011). Table 1 provides an overview of effect size measures that can be used for direct and indirect effects and that meet the basic criteria to be deemed useful as an indicator of the size of an effect (Preacher & Kelley, 2011; Wen & Fan, 2015). Specifically, each of these measures quantifies the size of the effect independently of the sample size and the unit of measurement of the variables. Each of these effect size measures is also zero when the unstandardized estimate of the effect it quantifies is zero, and each satisfies the requirement of being a monotonic function of the unstandardized estimate, a criterion several effect sizes of indirect effects lack, including κ^2 (Preacher & Kelley, 2011) and the ratio of *ab* to *c* (Wen & Fan, 2015). These effect sizes do not require raw data and can be estimated for path and latent variable models. Confidence interval limits can also be computed for each using bootstrapping.

Effect size measures for direct effects. For the *b*-paths and c', we evaluate three effect size measures and present equations to convert them (see Table 2).[1] A simple effect size measure is the standardized estimate, which is

$$b_s = b\frac{SD_X}{SD_Y}, \tag{4}$$

where b is the unstandardized point estimate of the effect from *X* on *Y* and *SD* is the standard deviation. This standardized effect is an estimate of how much *Y* is expected to increase or decrease in standard deviations

Table 2. Equations for Converting Effect Sizes

	As a function of		
Calculating	β_j	$r^2_{sp_j}$	Cohen's f^2
β_j	—	$\sqrt{\dfrac{r^2_{sp_j}}{1-R^2_j}}$	$\sqrt{\dfrac{f^2(1-R^2)}{1-R^2_j}}$
$r^2_{sp_j}$	$\beta^2_j(1-R^2_j)$	—	$f^2(1-R^2)$
Cohen's f^2	$\beta^2_j\dfrac{1-R^2_j}{1-R^2}$	$\dfrac{r^2_{sp_j}}{1-R^2}$	—

Note: β_j = standardized estimate of the *j*th predictor variable; $r^2_{sp_j}$ = squared semipartial correlation between the *j*th predictor variable and the outcome; R^2 = proportion of explained variance in the outcome; R^2_j = proportion of explained variance in the *j*th predictor variable by the other predictor variables; the expression $1-R^2_j$ is known as the tolerance.

if X increases by 1 *SD*. If the predictor variable, X, is dichotomous, it is common to do a partial standardization by dividing the unstandardized estimate by the standard deviation of the dependent variable, Y, because the metric of the dichotomous variable is meaningful:

$$b_{ps} = \frac{b}{SD_Y}. \tag{5}$$

Standardized effects can easily be interpreted and calculated for simple and complex mediation models that include covariates, multiple predictor variables and mediators, and latent variables. They also provide information about the direction of the association between the variables. However, they also have several limitations. One is that standardized effects depend on both the proportion of total variance explained in the outcome and the proportion of variance explained by the other predictor variables in the model. This dependency limits the comparability of standardized effects across different models and studies. This limitation is related to another one, which is the lack of widely accepted benchmarks for classifying these effects as small, medium, or large. Another limitation is that standardized effects can be larger than 1 and smaller than −1 (e.g., Jöreskog, 1999), rendering their interpretation less intuitive than that of other effect sizes, such as the squared semipartial correlation. Finally, standardized effects are inappropriate for assessing the size of interaction and nonlinear effects because the standard deviations of product terms lack meaningful interpretation.

Another effect size measure is the squared semipartial correlation, r^2_{sp}. For c' in Model B, the squared semipartial correlation, $r^2_{Y(X.M_1M_2)}$, is the squared correlation between Y and X after partialling out the effects of the mediators (and any other variables linked to Y) from the variable X. The squared semipartial correlation for c' is the proportion of variance in Y that is uniquely explained by X, controlling the latter for M_1 and M_2. It is also the difference in the proportion of variance explained between the full mediation model, which includes all variables, and the reduced model, which does not include X:

$$r^2_{Y(X.M_1M_2)} = R^2_{Y.XM_1M_2} - R^2_{Y.M_1M_2}. \tag{6}$$

The squared semipartial correlation can also be calculated as a function of the standardized effect (Cohen & Cohen, 1975):

$$r^2_{Y(X.M_1M_2)} = c'^2_s(1 - R^2_{X.M_1M_2}), \tag{7}$$

where $R^2_{X.M_1M_2}$ refers to the proportion of variance explained in X by the other predictor variables, M_1 and M_2.[2] The semipartial correlation, sometimes called "part correlation," can be estimated for models with observed and latent variables using SEM techniques (see Preacher, 2006). For Model B, Figure A1 in the Appendix shows a path diagram estimating the semipartial correlation for c', $r_{Y(X.M_1M_2)}$. The squared semipartial correlation is a standardized measure bounded by 0 and 1. Although there are no conventions for squared semipartial correlations that allow an interpretation of the size of an effect as small, medium, or large, any effect that uniquely explains 1% or more of the total variance can be considered substantial.

An interesting and often used effect size measure for specific effects in models with multiple predictors is Cohen's f^2. For c', Cohen's f^2 is:

$$f^2_{c'} = \frac{r^2_{Y(X.M_1M_2)}}{1 - R^2_Y}, \tag{8}$$

where R^2_Y is the proportion of variance explained in Y by all variables that directly predict Y. Cohen's f^2 is a signal-to-noise ratio that quantifies the proportion of variance uniquely explained by X relative to the proportion of variance that is not explained. The fact that f^2 depends on the proportion of variance not explained is an interesting characteristic and distinguishes Cohen's f^2 from the effect size measures discussed above. It is an effect size measure commonly used in power analysis for regression models such that $f^2 = 0.02$, 0.15, and 0.35 represent small, medium, and large effect sizes, respectively. Although these benchmarks, like any convention, are somewhat arbitrary (Cohen, 1988), they are widely accepted and allow for a more nuanced interpretation of the results. Other advantages of Cohen's f^2 are that it can be used in models with multiple predictors and covariates and for assessing the size of interaction and nonlinear effects (e.g., Smithson & Shou, 2017). It can also be calculated when using implicit-mediation analysis, which has been proposed to assess causal effects in treatment designs (Bullock & Green, 2021; Gerber & Green, 2012). Disadvantages are that Cohen's f^2 can range from 0 to infinity and that its interpretation is not as intuitive as that of other effect size measures.

Effect size measures for indirect effects. For the indirect effects, we focus on three effect size measures (see Table 1). A simple effect size measure is the standardization of *ab*, which is the unstandardized point estimate multiplied by the ratio of the *SD* of *X* to the *SD* of *Y* (Alwin & Hauser, 1975):

$$ab_s = ab\frac{SD_X}{SD_Y}. \tag{9}$$

This standardized indirect effect is an estimate of how much *Y* is expected to increase or decrease in standard deviations if *X* increases by 1 *SD*.

If the *X* variable is dichotomous, MacKinnon (2008) recommended a partial standardization by standardizing the indirect effect only by the metric of *Y* because the metric of *X* is naturally meaningful:

$$ab_{ps} = \frac{ab}{SD_Y}. \tag{10}$$

This standardized indirect effect is an estimate of how much *Y* is expected to differ in standard deviations between the two groups. These standardizations of the indirect effect, ab_s and ab_{ps}, can be employed to specific and total indirect effects. As for direct effects, standardized indirect effects are not bounded, and there are no conventions for what effect estimate can be considered negligible, small, medium, or large.

Lachowicz et al. (2018) proposed parameter upsilon (υ) for simple mediation models, which reflects the variance in the outcome explained jointly by the mediator and the predictor variable, correcting for the spurious correlation associated with the indirect effect:

$$\upsilon = b_{b_s}^2 - (R_Y^2 - r_{XY}^2), \tag{11}$$

where b_{b_s} is the standardized estimate of the *b*-path in the simple model and r_{XY}^2 is the squared correlation between *X* and *Y*. This effect size addresses many limitations of previous effect-size estimates (e.g., Preacher & Kelley, 2011; Wen & Fan, 2015). However, the application of υ is limited to mediation models with one predictor variable and one mediator and no covariates. Other limitations are that υ can be smaller than 0, which is more likely to occur when there is suppression, and that it cannot be interpreted as a proportion of variance explained. There are also no benchmarks for classifying υ as small, medium, or large.

Conclusion. Undoubtedly, the r_{sp}^2 and Cohen's f^2 have advantages that other effect size measures do not have, including guidelines for interpreting effects as negligible, small, medium, or large in size or the quantification of the size of interaction and nonlinear effects. However, both these measures are limited to direct effects, and their calculation can be challenging if models are complex, such as those involving cross-lagged effects and three or more waves (e.g., Cole & Maxwell, 2003; Lucas, 2023). The effect size measures for indirect effects lack benchmarks for what can be deemed a small, medium, or large effect, a limitation they share with the standardized direct effect estimates. However, note that the benchmarks commonly used for *r*, Cohen's *d*, and Cohen's f^2 are inconsistent, which means that the classification of the size of the effect depends on the effect size measure (Correll et al., 2020).

For simple mediation models with no covariates, we recommend reporting *r* for the *a*-path, r_{sp}^2 or Cohen's f^2 for the *b*-path, and υ for the indirect effect. For complex mediation models in which υ cannot be calculated, we suggest reporting standardized indirect effects in addition to r_{sp}^2 or Cohen's f^2 for the direct effects. Regardless of the effect size, researchers are encouraged to discuss the practical implications of finding an effect of that size. In addition, we recommend reporting the unstandardized estimates of all effects in a model because they can be practically meaningful and important for certain interpretations.

Being cautious in claiming full mediation

The second recommendation we have concerns the distinction between partial versus full mediation. Although this notion is appealing to many researchers because it facilitates the interpretation and discussion of findings of a mediation study, its testing, particularly full mediation, raises several questions, as discussed above and by others (e.g., Hayes, 2018; Montoya & Hayes, 2017; Preacher & Kelley, 2011; Rucker et al., 2011; Wood et al., 2008). We believe that the distinction of partial and full mediation makes conceptual sense and suggest retaining it for theoretical models. Claiming full mediation based solely on statistical results can be problematic because it requires the elimination of alternative explanations. In contrast, claiming partial mediation is less controversial, if at all, for two reasons. First, in a model with multiple simultaneous mediators, each mediator alone partially mediates the effect of the cause on the outcome. In a simple mediation model, it is unlikely that there would

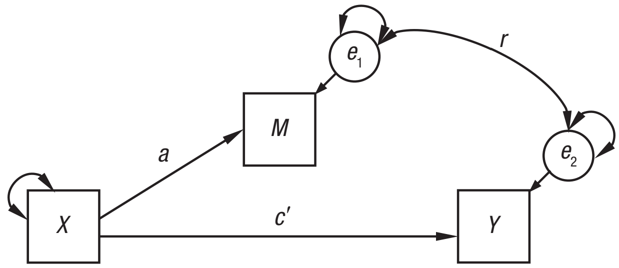

Fig. 2. Path diagram of a simple mediation model with a residual covariance.

be an indirect effect but no c' effect. That said, partial mediation is what a researcher can expect to find when testing mediation in a simple or complex model with multiple mediators. Second, the term "partial effect" is used in multiple regression analysis to refer to the change in an outcome for every unit change in a predictor variable holding the other predictor variables constant.

Sensitivity analysis

Sensitivity analysis can be employed to evaluate the robustness of the *b*-path to omitted confounders (Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010). To determine how large the effect of an omitted confounder that affects both the mediator, *M*, and the outcome, *Y*, would need to be for *b* to be 0 (or another value of interest), the correlation, *r*, between the residuals of *M* and *Y* can be calculated (Imai, Keele, & Yamamoto, 2010). The stronger this correlation is, the more robust the *b*-path is to omitted confounders that affect both *M* and *Y*. The residual correlation can be estimated using SEM or the R package *mediation* (Tingley et al., 2014). Using SEM, researchers can estimate this correlation by fixing *b* to 0 or another meaningful value and adding a covariance between the residuals of *M* and *Y* (see Fig. 2). A crucial question that remains is at what point a residual correlation can be considered sufficiently large enough to claim an effect is robust to violation of omitted confounders.

Power considerations

When determining the sample size for a planned study, it is essential that all effects of a mediation model that are expected to be substantial in magnitude are included in the power analysis. If c' is expected to be substantial, it is necessary to demonstrate that the study has sufficient power to detect a substantial c' effect in addition to the indirect effect and its constituents because if power is low, the likelihood that a statistically significant effect reflects the true effect is reduced (Button et al., 2013). Note that the *b*-paths are often smaller in size than the *a*-paths because the *b*-paths are partial effects and the power to detect a *b*-path can decrease as the *a*-path increases (Fritz et al., 2012), which is particularly problematic when *X* is a manipulated variable but *M* is not (MacKinnon, 2008).

Conducting a power analysis after data collection, in which power is estimated based on the sample size, effect size, and significance level, has its proponents (e.g., Arend & Schäfer, 2019; Mathieu et al., 2012; Onwuegbuzie & Leech, 2004). Indeed, observed power can be useful for researchers interested in determining the sample size needed for a subsequent study examining the same or similar variables (e.g., K.-H. Yuan & Maxwell, 2005). In addition, post hoc power analysis can provide insight into the power of different effects in a mediation model. Using bootstrapping techniques, power estimates can also reveal two rare but possible scenarios: a high *p* value with high power (> .80) or a low *p* value with low power (e.g., < .50). In either case, the data should be examined further, particularly for potential outliers and violations of the assumptions underlying the statistical test used. What we do not recommend is the use of post hoc power analysis to explain away nonsignificant effects, which could be the result of a small sample size, an effect that is negligible in size, measurement error, unmeasured confounders, or the use of an inappropriate statistical method (see also Wang & Rhemtulla, 2021). Clearly, it cannot be emphasized enough that power estimates from an actual study should not be used to interpret results (e.g., Giner-Sorolla et al., 2024; Hoenig & Heisey, 2001; Pek et al., 2024).[3]

Possible interaction effects

If *X* is a randomized variable, MacKinnon et al. (2020) recommend assessing the effect of the *XM* interaction on *Y*. This test provides a check of the assumption that b and c' do not differ across conditions. The effect of the *XM* interaction on *Y* can be estimated by adding the product of *X* and *M* to the simple mediation model:

$$Y = i_Y + bM + c'X + hXM + e_Y. \tag{12}$$

The effect of *X* on *M* is given by

$$M = i_M + aX + e_M. \tag{13}$$

For a binary variable, where 0 represents the control group and 1 represents the treatment group, five effects are of particular interest (as in MacKinnon et al., 2020): The mediating effect under the control condition is equal ab and referred to as the pure natural indirect effect. Under the treatment condition, this effect is equal $a(b+h)$ and called the total natural indirect effect. The direct effect of *X* on *Y* for the control condition is equal

$c' + bi_M$ and called the pure natural direct effect. For the treatment condition, this effect is equal $c' + b(i_M + a)$ and referred to as the total natural direct effect. The total effect of X on Y is equal c.

Illustrations

We illustrate the assessment of mediation for hypothetical and publicly available data. The first illustration uses a variance-covariance matrix for three variables with various sample sizes. The second illustration uses longitudinal data. We used R (R Core Team, 2024) and the package *lavaan* (Rosseel, 2012) for the analyses. Sensitivity analysis was conducted to assess the robustness of the b-paths to omitted confounders by calculating the correlation between the residuals of the corresponding mediator and outcome for b = 0. Percentile bootstrap confidence intervals (CIs) were calculated for the effects. For the b-paths and c', Cohen's f^2 was calculated using Equation 7. Post hoc power simulations were conducted adopting the R code developed by Ledermann et al. (2022) for the mediation actor–partner interdependence model (Ledermann et al., 2011). This code estimates the power for the direct, indirect, and total effects, as well as the differences between effects using the delta method (Sobel, 1982). Although this approach is practical, it tends to underestimate the observed power for detecting indirect effects, especially compared with the bootstrap method. The code for R to run the analyses can be accessed at https://github.com/thomasledermann/MediationEffectSize.

Hypothetical data

Data and statistical analyses. We estimated a simple mediation model using a covariance matrix as input data. We followed MacKinnon et al. (2002) and fixed a to 0.14 and b and c' to 0.39 to obtain the implied variance-covariance matrix, which served as input data. The sample sizes used were 50, 150, and 250. The squared semipartial correlation was calculated using an adapted version of the model shown in Figure A1 for one mediator. Percentile CIs were calculated using parametric bootstrapping with 10,000 bootstrap samples and the function *mvrnorm* from the package *MASS* (Venables & Ripley, 2002). For estimating power, the correlations between the three variables were calculated ($r_{xy} = .140$, $r_{my} = r_{xy} = .427$) to run a Monte Carlo simulation with 10,000 random samples for each sample size (for details, see Ledermann et al., 2022).

Results. Table 3 provides the results. The direct effects were small in size, as indicated by the standardized estimates for a and Cohen's f^2 for b and c', and the total effect was medium in size (standardized estimate > 0.30). The standardized estimate of the indirect effect was 0.052, and υ was .003. The residual correlation between the mediator and the outcome was .382 (N = 50: 95% CI = [.154, .625]; N = 150: 95% CI = [.271, .537]; N = 250: 95% CI = [.305, .511]), indicating a substantial association. For N = 50, the b-path, c', and total effect were statistically significant; c' was significantly stronger in size than the nonsignificant mediating effect. This indicates that X is only directly related. The same pattern emerged for N = 150, leading to the same conclusion. For N = 250, all effects were significant, suggesting partial mediation, and the direct effect c' was significantly stronger than the mediating effect. The proportions of the variance explained were 2% for the mediator and 32% for the outcome. Power substantially increased as the sample size increased. Although all effects were significant for 250 cases, power was substantially lower for the a-path and the mediating effect compared with the other effects.

Longitudinal data

We used longitudinal data from the Midlife in the United States (MIDUS) survey (Ryff et al., 2007). We merged the data sets MIDUS 2 Project 1 and Biomarker Project (2004–2005) and MIDUS 3 Project 1 (2013), which are available from the Inter-University Consortium for Political and Social Research.

Sample and measures. There were 945 adults participating in the study (age: M = 52.22 years, SD = 9.63). We used emotional abuse in childhood as predictor variable (Childhood Trauma Questionnaire; 21% or 198 abused) and matched participants not reporting emotional abuse on their age and gender with the abused participants using the function *matchControls* from the R package *e1071* (Meyer et al., 2021). This resulted in 198 abused and 198 not-abused participants (396 total) who provided complete data on self-esteem (Rosenberg Self-Esteem Scale; Rosenberg, 1965) and negative affect from the Positive and Negative Affect Schedule (Watson et al., 1988) measured in MIDUS 2 (Time 1) and MIDUS 3 (Time 2). Table 4 provides the descriptive statistics and the correlations of these variables.

Mediation model and statistical analyses. We estimated a longitudinal mediation model to test the hypothesis that self-esteem at Time 1 (M) mediates the effect of childhood abuse (X) on negative affect at Time 2 (Y). Following common recommendations (e.g., Maxwell & Cole, 2007; Mitchell & Maxwell, 2013), we included negative affect at Time 1 as a predictor variable to estimate and control for its autoregressive (stability) effect. We also estimated the effect between childhood abuse and negative affect at Time 1, which has the advantage that the statistical

Table 3. Results of the Analysis of the Simple Mediation Model

N	Effect	Estimate [95% CI]	SE	z	p	Standardized estimate [95% CI]	υ [95% CI]	r_{sp}^2 [95% CI]	f^2 [95% CI]	Power
50										
	$X \rightarrow M$ (a)	0.140 [−0.140, 0.424]	0.140	1.000	.317	0.140 [−0.130, 0.398]				0.186
	$M \rightarrow Y$ (b)	0.390 [0.140, 0.630]	0.123	3.178	.001	0.374 [0.134, 0.580]		.137 [.017, .331]	0.202 [0.023, 0.645]	0.871
	$X \rightarrow Y$ (c')	0.390 [0.142, 0.641]	0.123	3.178	.001	0.374 [0.137, 0.583]		.137 [.017, .324]	0.202 [0.022, 0.624]	0.868
	ab	0.055 [−0.053, 0.189]	0.057	0.954	.340	0.052 [−0.052, 0.174]	.003 [.000, .028]			0.070
	$ab + c'$	0.445 [0.197, 0.711]	0.133	3.337	.001	0.427 [0.176, 0.633]				0.902
	$ab - c'$	−0.335 [−0.607, −0.033]	0.138	−2.438	.015	−0.322 [−0.569, −0.029]				0.659
150										
	$X \rightarrow M$ (a)	0.140 [−0.017, 0.301]	0.081	1.732	.083	0.140 [−0.017, 0.295]				0.420
	$M \rightarrow Y$ (b)	0.390 [0.244, 0.530]	0.071	5.504	< .001	0.374 [0.238, 0.498]		.137 [.057, .237]	0.202 [0.080, 0.394]	1.000
	$X \rightarrow Y$ (c')	0.390 [0.253, 0.529]	0.071	5.504	< .001	0.374 [0.243, 0.493]		.137 [.056, .242]	0.202 [0.077, 0.409]	1.000
	ab	0.055 [−0.006, 0.126]	0.033	1.652	.099	0.052 [−0.006, 0.116]	.003 [.000, .013]			0.357
	$ab + c'$	0.445 [0.294, 0.596]	0.077	5.780	< .001	0.427 [0.289, 0.548]				1.000
	$ab - c'$	−0.335 [−0.484, −0.177]	0.079	−4.222	< .001	−0.322 [−0.459, −0.171]				0.976
250										
	$X \rightarrow M$ (a)	0.140 [0.018, 0.260]	0.063	2.236	.025	0.140 [0.018, 0.257]				0.615
	$M \rightarrow Y$ (b)	0.390 [0.278, 0.498]	0.055	7.105	< .001	0.374 [0.272, 0.467]		.137 [.074, .216]	0.202 [.102, .347]	1.000
	$X \rightarrow Y$ (c')	0.390 [0.281, 0.501]	0.055	7.105	< .001	0.374 [0.276, 0.471]		.137 [.072, .216]	0.202 [.099, .347]	1.000
	ab	0.055 [0.007, 0.107]	0.026	2.133	.033	0.052 [0.007, 0.100]	.003 [.000, .010]			0.580
	$ab + c'$	0.445 [0.326, 0.562]	0.060	7.462	< .001	0.427 [0.322, 0.525]				1.000
	$ab - c'$	−0.335 [−0.458, −0.213]	0.062	−5.451	< .001	−0.322 [−0.437, −0.202]				0.999

Note: Five thousand bootstrap samples were used to calculate the percentile bootstrap confidence intervals. r_{pb} = point biserial correlation; r_{sp}^2 squared semipartial correlation; f^2 = Cohen's f^2; CI = confidence interval.

Table 4. Descriptive Statistics and Correlations

Variable	1	2	3	4
1. Childhood abuse	—			
2. Self-esteem	−.261	—		
3. Negative affect Time 1	.218	−.581	—	
4. Negative affect Time 2	.204	.442	.502	—
M	0.000	36.914	1.633	1.594
SD	1.001	7.291	0.540	0.542
Minimum	−1.000	11.000	1.000	1.000
Maximum	1.000	49.000	4.000	4.400
Skewness	0.000	−0.737	1.305	1.620
Kurtosis	−2.010	0.196	2.065	4.322

Note: $N = 396$. Emotional childhood abuse was coded −1 = not emotionally abused and 1 = emotionally abused; 50% reported emotional abuse.

model is just identified ($df = 0$), meaning that it perfectly reproduces the sample covariance matrix. This setup enables the estimation of the total effect, which is identical to the effect between childhood abuse and negative affect at Time 2 without the mediator and negative affect at Time 1. Although the structure of this saturated model is identical to the mediation Model B of Figure 1, the focus in this longitudinal model is on the simple indirect effect through self-esteem and the direct effect and total effect between childhood abuse and negative affect at Time 2. We used effect coding for childhood abuse such that −1 = no abuse and 1 = emotional abuse in childhood. We calculated point biserial correlations for the effects from the predictor variable to self-esteem and negative affect. The squared semipartial correlation was calculated for the autoregressive effect, the effect from self-esteem to negative affect, and c' using Equation 6. The effects were standardized using Equations 4 (b-paths), 5 (a-paths and c'), and 10 (indirect effects and total effect). Percentile bootstrap CIs were calculated using 5,000 bootstrap samples (Fossum & Montoya, 2023). Bootstrapping with 10,000 bootstrap samples was used to estimate power for each effect.

Results. The results of the mediation analysis are presented in Table 5. Figure 3 shows the path diagram of the model. Both direct effects that make up the indirect effect through self-esteem (a_1 and b_1) were negative and statistically significant. Cohen's f^2 indicates that the effect between childhood abuse and self-esteem (a_1) was medium in size and that the effect between self-esteem and negative affect at Time 2 (b_1) was small. The indirect effect through self-esteem (a_1b_1) was also significant, indicating that the effect of childhood abuse on negative affect at Time 2 was transmitted through individuals' self-esteem. The direct effect c' was negligible in size and not statistically significant, suggesting that childhood abuse did not have a direct effect on negative affect at Time 2 above and beyond the effects of negative affect at Time 1 and self-esteem. Looking at the direct effects to and from negative affect at Time 1, we found that both effects, a_2 and b_2, were positive, statistically significant, and small in size. The total effect was also significant, and it was significantly stronger than the indirect effect through self-esteem and the direct effect c', which contributed 27.3% and 33.6% to the total effect, respectively. These results on the indirect effects (a_1b_1 and a_2b_2) suggest that self-esteem only partially mediated the effect of childhood abuse on negative affect at Time 2. No significant difference emerged between the indirect effect through self-esteem (a_1b_1) and the direct effect c'. The proportions of variance explained in self-esteem and negative affect at Times 1 and 2 were 6.8%, 4.7%, and 29.0%, respectively.

The residual correlation between self-esteem and negative affect at Time 2 was −.165 (95% CI = [−.242, −.080]), indicating a substantial association, especially considering the inclusion of the autoregressive effect. Table 5 also presents the power estimates from the bootstrap power simulations. These estimates ranged from 0.960 to 0.999 for a_1, b_1, a_1b_1, and the total effect. Changing the sample size, power simulations revealed that 262 individuals would have been needed to achieve a power of at least 0.80 for these effects (0.804 for the indirect effect through self-esteem).

Conclusion. The results indicate that self-esteem partially mediated the effect of emotional childhood abuse on negative affect at Time 2, contributing 27.3% to the total effect. The power estimates reveal that this study was well powered to detect the substantial direct and indirect effects and that a sample size of 262 would have been sufficient to achieve a power of 0.80.

Table 5. Results for Emotional Childhood Abuse Predicting Negative Affect at Time 2 Through Self-Esteem and Negative Affect at Time 1

Effect	Estimate [95% CI]	SE	$\|z\|$	p	Standardized estimate [95% CI]	r_{pb} [95% CI]	r_{sp}^2 [95% CI]	f^2 [95% CI]	Power
Direct effects									
A → S (a_1)	−1.804 [−2.602, −1.221]	0.353	5.391	< .001	−0.261 [−0.350, −0.124]	−.327 [−.438., −.214]			.999
A → N1 (a_2)	0.117 [0.067, 0.170]	0.026	4.440	< .001	0.218 [0.124, 0.360]	.273 [.157, .381]			0.995
S → N2 (b_1)	−0.016 [−0.025, −0.007]	0.004	4.024	< .001	−0.212 [−0.317, −0.102]		.029 [.007, .064]	0.041 [0.010, 0.094]	0.973
N1 → N2 (b_2)	0.365 [0.226, 0.497]	0.052	6.969	< .001	0.364 [0.216, 0.502]		.087 [.034, .165]	0.123 [0.044, 0.268]	1.00
A → N2 (c')	0.037 [−0.008, 0.084]	0.024	1.563	.118	0.069 [−0.016, 0.151]		.004 [.000, .021]	0.006 [0.000, 0.031]	0.352
Indirect and total effects									
a_1b_1	0.030 [0.012, 0.053]	0.009	3.225	.001	0.056 [0.023, 0.094]				0.960
a_2b_2	0.043 [0.020, 0.072]	0.011	3.745	< .001	0.079 [0.036, 0.133]				0.991
$a_1b_1 + a_2b_2$	0.073 [0.042, 0.110]	0.015	4.843	< .001	0.135 [0.079, 0.194]				0.999
Total effect	0.110 [0.058, 0.163]	0.027	4.137	< .001	0.204 [0.108, 0.293]				0.986
Contrasts									
$a_1b_1 - c'$	−0.007 [−0.056, 0.045]	0.027	0.268	.788					0.054
Total effect − c'	0.073 [0.042, 0.110]	0.015	4.843	< .001					0.999
Total effect − a_1b_1	0.080 [0.031, 0.129]	0.026	3.086	.002					0.880

Note: Partial standardization was used for a_1, a_2, c', a_1b_1, a_2b_2, $a_1b_1 + a_2b_2$, and $a_1b_1 + a_2b_2 + c'$. Five thousand bootstrap samples were used to calculate the percentile bootstrap confidence intervals. CI = confidence interval; r_{pb} = point biserial correlation; r_{sp}^2 squared semipartial correlation; f^2 = Cohen's f^2; A = emotional childhood abuse (–1 = not emotionally abused, 1 = emotionally abused); S = self-esteem; N1 and N2 = negative affect at Time 1 and Time 2, respectively. Total effect = $a_1b_1 + a_2b_2 + c'$.

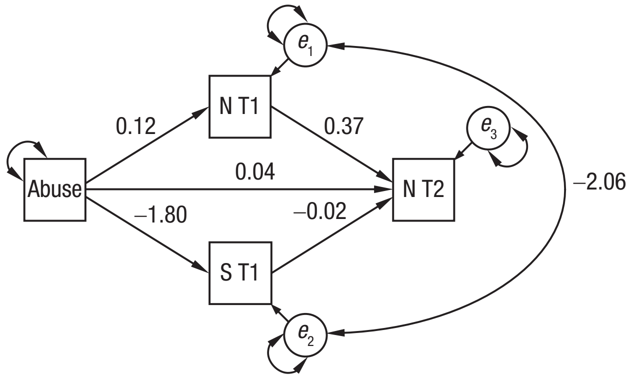

Fig. 3. Path diagram of the mediation model with child abuse as the predictor variable, self-esteem as mediator, and negative affect as outcome. Abuse = Child Abuse (−1 = no abuse, 1 = emotional abuse), N = Negative Affect, S = Self-Esteem, T1 = time 1, T2 = time 2. The numbers represent unstandardized estimates.

Discussion

The development of sophisticated frameworks and theories of causal relationships plays a pivotal role in advancing the knowledge base and is an indicator of the maturation of a discipline. Theoretical models involving multiple intervening variables often posit partial or full mediation (e.g., Karremans et al., 2017; Randall & Bodenmann, 2009; Schmader & Sedikides, 2018). From a theoretical perspective, we believe that this notion of partial versus full mediation makes sense and should continue to be used because it enhances theoretical precision and theoretical understanding of the mechanism through which a cause brings about an effect. Thus, theoretical frameworks involving one or multiple mediators should not only include all theoretically relevant variables and prescribe which variable comes first and what is the relationship between the variables (see also Sutton & Staw, 1995) but also be clear on whether a cause is expected to influence an outcome directly and indirectly or only indirectly. If partial or full mediation is equally plausible in a theoretical model, a statement like the one made by Rusbult et al. (1998) that the mediator "partially or wholly mediates the effects of" (p. 383) the predictor variables on the outcome is useful for scholars using a theoretical model to guide their research and for practitioners making decisions about where it is appropriate to intervene.

From a statistical perspective, the testing of the distinction between partial versus full mediation has been found to be problematic (e.g., Hayes, 2018; Montoya & Hayes, 2017; Preacher & Kelley, 2011; Rucker et al., 2011; Wood et al., 2008). We have extended this critique by underscoring that a claim of full statistical mediation is a claim of a null result, necessitating the elimination of alternative explanations for the absence of an effect, such as low power or poor measures. Less problematic are claims of partial mediation. In fact, a researcher can expect to find partial mediation in both a simple model with a single mediator and no covariates and complex models with multiple mediators in which each mediator alone may partially mediate the effect of a cause on an outcome.

Exceptions to this expectation are noteworthy. One is a longitudinal mediation model with cross-lagged and autoregressive effects involving three or more waves, where researchers often start with a model that does not include the c'-paths for the autoregressive effects (Cole & Maxwell, 2003; Jose, 2016; Maxwell & Cole, 2007; Mitchell & Maxwell, 2013; Zhang et al., 2018). A second situation is a mediation model with an instrumental variable for the mediator where the instrument affects *M* but not *Y* (e.g., Kline, 2015; MacKinnon & Pirlott, 2015; Sobel, 2008). Randomized binary variables serve as ideal instruments for the mediator when the randomized variable has a direct effect on *M* but not on *Y*. Instrumental variables are also required in nonrecursive models with reciprocal effects. An example is Kenny's (1996) mutual influence model (see also Ledermann & Kenny, 2017), which has been designed to assess reciprocal effects between two partners' outcomes. In this model, each instrumental variable affects one's own outcome but not the partner's outcome (e.g., reciprocal effects between partners' behavior and their attitudes as instrumental variables).

In assessing mediation mechanisms, we echo Kline (2015) in advocating against an overreliance on null hypothesis significance testing as a decision rule and, along with others (e.g., Lee et al., 2021), recommend conducting additional analyses. One recommendation is the calculation of effect sizes for the direct and indirect effects. The effect size measures for direct effects, especially the squared semipartial correlation and Cohen's f^2, possess desired characteristics that current effect size measures for indirect effects do not have, such as benchmarks for classifying effects as negligible, small, medium, or large and a broad applicability in simple and complex models, including models with interaction and nonlinear effects. Cohen's f^2 in particular is a popular effect size measure frequently used in power analysis for regression models with multiple predictor variables that can facilitate the interpretation and communication of findings. We generally recommend the practice of reporting and interpreting effect sizes whenever possible because they are crucial for designing future studies with adequate power and aggregating results across studies. What remains unknown is how accurate the percentile bootstrap CIs are for the effect size measures discussed in this article, and as others have noted (e.g., Lachowicz et al., 2018), other CIs, such as bias-corrected CIs, may be superior.

Sensitivity analysis can provide insights into the robustness of the effects to potential unmeasured confounders (Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010). The correlation between the residuals of the mediator and the outcome is a standardized measure of the effect of an omitted confounder affecting both the mediator and the outcome. Although this approach is straightforward for *b*-paths, the analysis becomes complicated if an omitted confounder confounds not only the mediator and outcome but also the cause (e.g., Smith & VanderWeele, 2019; Tofighi, 2021; VanderWeele, 2010). For randomized designs, the testing of the *XM* interaction has been recommended (MacKinnon et al., 2020). This approach provides insights into whether the effects in a mediation model differ across conditions.

Power analysis is crucial in the planning phase of a study and can provide insight into the actual power of specifics effects. When determining an appropriate sample size for a planned study, it is essential to consider all effects that are expected to be substantial in size. Power estimates, especially from simulations, are a worthwhile afterthought because they can provide insights into the power of different effects and whether a smaller sample size would have been sufficient, which can be valuable information for both researchers and resource providers (see also K.-H. Yuan & Maxwell, 2005). However, it is crucial to reiterate that power estimates should not be used to interpret the results of a study (e.g., Hoenig & Heisey, 2001; Pek et al., 2024).

Although the consideration of the effect sizes, along with sensitivity analysis and power considerations, provides a more complete picture of the mechanism by which a cause impacts an outcome, a couple of points are important to note. First, no statistical effect is an unbiased estimate of a true effect because of the omission of putative confounders or mediators or violations of the assumptions underlying the mediation analysis (e.g., MacKinnon, 2008; MacKinnon & Pirlott, 2015; Pearl, 2014; VanderWeele, 2015). Moreover, it is important to keep in mind that all models, whether theoretical or statistical, are an approximation of the reality and that any direct effect may itself be mediated, highlighting the need to focus on not only the macroscopic role of mediators but also their microscopic role in more fine-grained models (MacKinnon, 2008). Second, the use of Bayesian analysis is becoming increasingly popular (Enders et al., 2013; Miočević et al., 2018; Y. Yuan & MacKinnon, 2009) for testing mediation in small samples. Finally, interpretability challenges can arise when the results are inconsistent. For example, c' may be statistically significant but negligible in size, which may be due to a large sample size, inflating the statistical power.

In conclusion, we have shown that the distinction between partial and full mediation is useful in the context of theoretical models because it provides common language that allows researchers to describe mediating mechanisms in simple terms many scientists are familiar with. Statistically, any claim of full mediation is essentially a claim of a null result, which requires the elimination of alternative explanations. To ensure a more nuanced understanding of mediation results, especially in underpowered or overpowered studies, we recommend accompanying unstandardized estimates with effect sizes, particularly for the direct effects, and the consideration of power. These recommendations extend beyond mediation analysis and are especially relevant to regression analysis and SEM. Conducting sensitivity analysis and testing for possible interaction effects in randomized designs can provide further insights into the mechanism through which a predictor variable affects an outcome.

Appendix

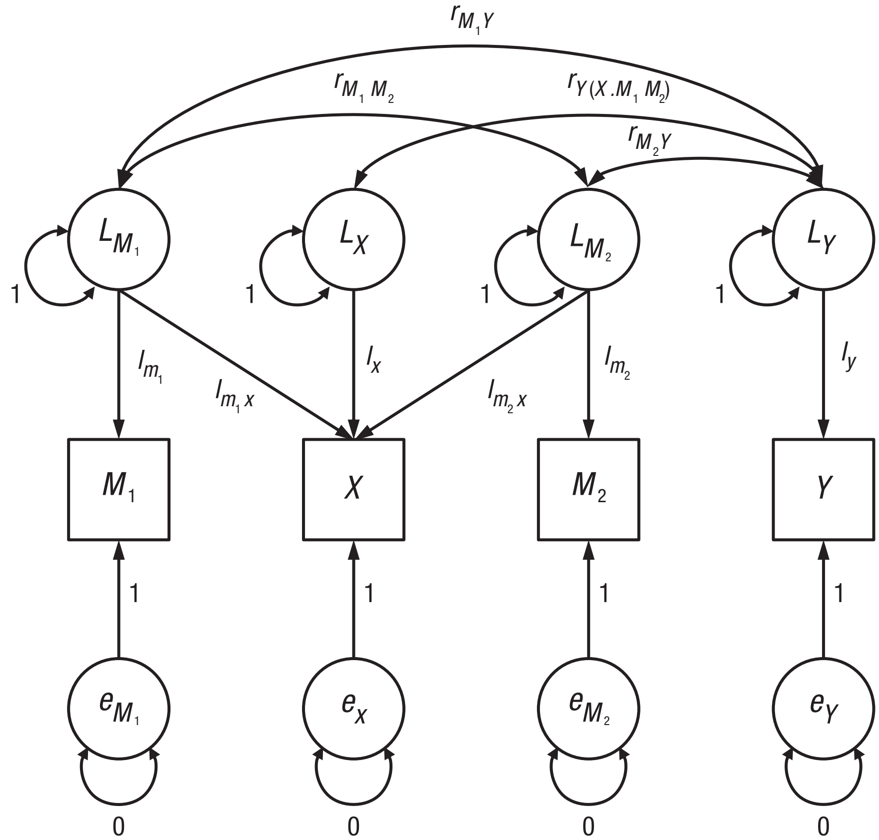

Fig. A1. Path diagram for estimating the semipartial correlation between X and Y controlling for M_1 and M_2.

Transparency

Action Editor: Pamela Davis-Kean
Editor: David A. Sbarra
Author Contributions

Thomas Ledermann: Conceptualization; Formal analysis; Methodology; Software; Writing – original draft.
Myriam Rudaz: Writing – review & editing.
Matthew Fritz: Methodology; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iDs

Thomas Ledermann https://orcid.org/0000-0002-4976-6942
Myriam Rudaz https://orcid.org/0000-0003-0550-3558
Matthew S. Fritz https://orcid.org/0000-0002-5885-6341

Acknowledgments

We thank Siegfried Macho and Damaris Aschwanden for their helpful feedback on an earlier version of this article and David Kenny for valuable discussion on the topic. We also thank two anonymous reviewers and the action editor for their helpful comments.

Notes

1. For the a-paths, a zero-order correlation can be computed for each path as an effect size measure if there are no covariates and if there is only one predictor variable. If there are covariates or multiple predictor variables, the effect size measures for the b-paths and c' can be used (i.e., b_s, b_{ps}, r^2_{sp}, Cohen's f^2).

2. The expression $(1 - R^2_{X.M_1M_2})$ is known as the tolerance, and the inverse of the tolerance is known as the variance inflation factor.
3. Hoenig and Heisey (2001) demonstrated that for data from a normal distribution, power is a monotonic function of p, meaning that power estimates provide no additional information for interpretating the results of a study beyond the p values.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, *40*(1), 37–47. https://doi.org/10.2307/2094445

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). https://doi.org/10.1037/0000165-000

Arend, M. G., & Schäfer, T. (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. *Psychological Methods*, *24*(1), 1–19. https://doi.org/10.1037/met0000195

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*(6), 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173

Berk, R. A. (2005). Randomized experiments as the bronze standard. *Journal of Experimental Criminology*, *1*(4), 417–433. https://doi.org/10.1007/s11292-005-3538-2

Bullock, J. G., & Green, D. P. (2021). The failings of conventional mediation analysis and a design-based alternative. *Advances in Methods and Practices in Psychological Science*, *4*(4). https://doi.org/10.1177/25152459211047227

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*(4), 550–558. https://doi.org/10.1037/a0018933

Button, K., Ioannidis, J., Mokrysz, C., Nosek, B., Flint, J., Robinson, E., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cheung, M. W. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, *41*(2), 425–438. https://doi.org/10.3758/BRM.41.2.425

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum.

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*(4), 558–577. https://doi.org/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*(2), 300–315. https://doi.org/10.1037/a0033805

Correll J., Mellinger C., McClelland G. H., & Judd C. M. (2020). Avoid Cohen's "small," "medium," and "large" for power analysis. *Trends in Cognitive Sciences*, *24*(3), 200–207. https://doi.org/10.1016/j.tics.2019.12.009

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Enders, C. K., Fairchild, A. J., & MacKinnon, D. P. (2013). A Bayesian approach for estimating mediation effects with missing data. *Multivariate Behavioral Research*, *48*(3), 340–369. https://doi.org/10.1080/00273171.2013.784862

Fossum, J. L., & Montoya, A. K. (2023). When to use different inferential methods for power analysis and data analysis for between-subjects mediation. *Advances in Methods and Practices in Psychological Science*, *6*(2). https://doi.org/10.1177/25152459231156

Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. *Multivariate Behavioral Research*, *51*(5), 681–697. https://doi.org/10.1080/00273171.2016.1224154

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Fritz, M. S., Taylor, A. B., & MacKinnon, D. P. (2012). Explanation of two anomalous results in statistical mediation analysis. *Multivariate Behavioral Research*, *47*(1), 61–87. https://doi.org/10.1080/00273171.2012.640596

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. W. W. Norton.

Giner-Sorolla, R., Montoya, A. K., Reifman, A., Carpenter, T., Lewis, N. A., Aberson, C. L., Bostyn, D. H., Conrique, B. G., Ng, B. W., Schoemann, A. M., & Soderberg, C. (2024). Power to detect what? Considerations for planning and evaluating sample size. *Personality and Social Psychology Review*, *28*(3), 276–301. https://doi.org/10.1177/10888683241228328

Goldberg, L., Elliot, D., Clarke, G. N., MacKinnon, D. P., Moe, E., Zoref, L., Green, C., Wolf, S. L., Greffrath, E., Miller, D. J., & Lapin, A. (1996). Effects of a multidimensional anabolic steroid prevention intervention: The Adolescents Training and Learning to Avoid Steroids (ATLAS) program. *JAMA*, *276*(19), 1555–1562. https://doi.org/10.1001/jama.1996.03540190027025

Gonzalez, O., & MacKinnon, D. P. (2021). The measurement of the mediator and its influence on statistical mediation conclusions. *Psychological Methods*, *26*(1), 1–17. https://doi.org/10.1037/met0000263

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based perspective* (2nd ed.). The Guilford Press.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470. https://doi.org/10.1111/bmsp.12028

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis.

The American Statistician, *55*(1), 19–24. https://doi.org/10.1198/000313001300339897
Hoyle, R. H., & Kenny, D. A. (1999). Statistical power and tests of mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 195–222). Sage.
Hyman, H. H. (1955). *Survey design and analysis*. The Free Press.
Iacobucci, D. (2008). *Mediation analysis*. Sage.
Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761
Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. https://doi.org/10.1214/10-STS321
Jacob, R. T., Doolittle, F., Kemple, J., & Somers, M.-A. (2019). A framework for learning from null results. *Educational Researcher*, *48*(9), 580–589. https://doi.org/10.3102/0013189X19891955
James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, *69*(2), 307–321. https://doi.org/10.1037/0021-9010.69.2.307
Jöreskog, K. G. (1999). *How large can a standardized coefficient be?* http://www.ssicentral.com/lisrel/column2.htm
Jose, P. E. (2016). The merits of using longitudinal mediation. *Educational Psychologist*, *51*(3–4), 331–341. https://doi.org/10.1080/00461520.2016.1207175
Judd, C. M., & Kenny, D. A. (1981a). *Estimating the effect of social interventions*. Cambridge University Press.
Judd, C. M., & Kenny, D. A. (1981b). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*(5), 602–619. https://doi.org/10.1177/0193841X8100500502
Karremans, J. C., Schellekens, M. P., & Kappen, G. (2017). Bridging the sciences of mindfulness and romantic relationships: A theoretical model and research agenda. *Personality and Social Psychology Review*, *21*(1), 29–49. https://doi.org/10.1177/1088868315615450
Kenny, D. A. (1979). *Correlation and Causality*. Wiley.
Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships*, *13*(2), 279–294. https://doi.org/10.1177/0265407596132007
Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science*, *25*(2), 334–339. https://doi.org/10.1177/0956797613502676
Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). McGraw-Hill.
Kline, R. B. (2015). The mediation myth. *Basic and Applied Social Psychology*, *37*(4), 202–213. https://doi.org/10.1080/01973533.2015.1049349
Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, *23*(2), 244–261. https://doi.org/10.1037/met0000165
Landis, R. S., James, L. R., Lance, C. E., Pierce, C. A., & Rogelberg, S. G. (2014). When is nothing something? Editorial for the null results special issue of *Journal of Business and Psychology*. *Journal of Business and Psychology*, *29*(2), 163–167. https://doi.org/10.1007/s10869-014-9347-8
Le, T. P., Iwamoto, D. K., & Ching, T. H. W. (2024). Associations between gendered racism, racial identity, and nicotine use among Asian American men. *Journal of Clinical Psychology*, *80*(7), 1582–1595. https://doi.org/10.1002/jclp.23681
Lee, H., Cashin, A. G., Lamb, S. E., Hopewell, S., Vansteelandt, S., VanderWeele, T. J., MacKinnon, D. P., Mansell, G., Collins, G. S., Golub, R. M., McAuley, J. H. AGReMA group, Localio, A. R., van Amelsvoort, L., Guallar, E., Rijnhart, J., Goldsmith, K., Fairchild, A. J., Lewis, C. C., . . . Henschke, N. (2021). A guideline for reporting mediation analyses of randomized trials and observational studies: The AGReMA statement. *JAMA*, *326*(11), 1045–1056. https://doi.org/10.1001/jama.2021.14075
Ledermann, T., & Kenny, D. A. (2017). Analyzing dyadic data with multilevel modeling versus structural equation modeling: A tale of two methods. *Journal of Family Psychology*, *31*(4), 44–452. https://doi.org/10.1037/fam0000290
Ledermann, T., & Macho, S. (2009). Mediation in dyadic data at the level of the dyads: A structural equation modeling approach. *Journal of Family Psychology*, *23*(5), 661–670. https://doi.org/10.1037/a0016197
Ledermann, T., & Macho, S. (2015). Assessing mediation in simple and complex models. In L. Rivera (Ed.), *Structural equation modeling (SEM): Concepts, applications and misconceptions* (pp. 69–101). Nova Science Publishers.
Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling*, *18*(4), 595–612. https://doi.org/10.1080/10705511.2011.607099
Ledermann, T., Rudaz, M., Wu, Q., & Cui, M. (2022). Determine power and sample size for the simple and mediation Actor–Partner Interdependence Model. *Family Relations*, *71*(4), 1452–1469. https://doi.org/10.1111/fare.12644.
Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology*, *101*(6), 1174–1188. https://doi.org/10.1037/a0024776
Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2007). Structural equation modeling of mediation and moderation with contextual factors. In T. D. Little & J. A. Bovaird (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 207–230). Erlbaum.
Loh, W. W., Moerkerke, B., Loeys, T., & Vansteelandt, S. (2022). Disentangling indirect effects through multiple mediators without assuming any causal structure among the mediators. *Psychological Methods*, *27*(6), 982–999. https://doi.org/10.1037/met0000314
Lucas, R. E. (2023). Why the cross-lagged panel model is almost never the right choice. *Advances in Methods and Practices in Psychological Science*, *6*(1). https://doi.org/10.1177/25152459231158378
MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening constructs. *Psychological Review*, *55*(2), 95–107. https://doi.org/10.1037/h0056029

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis. Multivariate applications series*. Taylor & Francis Group/Lawrence Erlbaum Associates.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

MacKinnon, D. P., Goldberg, L., Clarke, G. N., Elliot, D. L., Cheong, J., Lapin, A., Moe, E. L., & Krull, J. L. (2001). Mediating mechanisms in a program to reduce intentions to use anabolic steroids and improve exercise self-efficacy and dietary behavior. *Prevention Science*, *2*(1), 15–28. https://doi.org/10.1023/a:1010082828000

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181. https://doi.org/10.1023/a:1026595011371

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*(1), 83–104. https://doi.org/10.1037/1082-989X.7.1.83

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*(1), 30–43. https://doi.org/10.1177/1088868314542878

MacKinnon, D. P., Valente, M. J., & Gonzalez, O. (2020). The correspondence between causal and traditional mediation analysis: The link is the mediator by treatment interaction. *Prevention Science*, *21*(2), 147–157. https://doi.org/10.1007/s11121-019-01076-4

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, *30*(1), 41–62. https://doi.org/10.1207/s15327906mbr3001_3

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*(5), 951–966. https://doi.org/10.1037/a0028380

Mathieu, J. E., & Taylor, S. R. (2006). Clarifying conditions and decision points for mediational type inferences in organizational behavior. *Journal of Organizational Behavior*, *27*(8), 1031–1056. https://doi.org/10.1002/job.406

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. https://doi.org/10.1037/1082-989X.12.1.23

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2021). *e1071: Misc functions of the department of statistics, probability theory group (formerly: E1071), TU Wien* [R Package Version 1.7-9]. Comprehensive R Archive Network (CRAN).

Miočević, M., Gonzalez, O., Valente, M. J., & MacKinnon, D. P. (2018). A tutorial in Bayesian potential outcomes mediation analysis. *Structural Equation Modeling*, *25*(1), 121–136. https://doi.org/10.1080/10705511.2017.1342541

Mitchell, M. A., & Maxwell, S. E. (2013). A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivariate Behavioral Research*, *48*(3), 301–339. https://doi.org/10.1080/00273171.2013.784696

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. https://doi.org/10.1037/met0000086

Muthén, B. O., Muthén, L. K., & Asparouhov, T. (2016). *Regression and mediation analysis using Mplus*. Muthén & Muthén.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Post hoc power: A concept whose time has come. *Understanding Statistics*, *3*(4), 201–230. https://doi.org/10.1207/s15328031us0304_1

Pearl, J. (2001). Direct and indirect effects. In *UAI'01: Proceedings of the Seventeenth Conference on Uncertainty in artificial Intelligence* (pp. 411–420). Association for Computing Machinery.

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

Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods*, *19*(4), 459–481. https://doi.org/10.1037/a0036434

Pek, J., Hoisington-Shaw, K. J., & Wegener, D. T. (2024). Uses of uncertain statistical power: Designing future studies, not evaluating completed studies. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000577

Preacher, K. J. (2006). Testing complex correlational hypotheses with structural equation models. *Structural Equation Modeling*, *13*(4), 520–543. https://doi.org/10.1207/s15328007sem1304_2

Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Sage. https://doi.org/10.4135/9781452272054.n2

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*(2), 93–115. https://doi.org/10.1037/a0022658

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Randall, A. K., & Bodenmann, G. (2009). The role of stress on close relationships and marital satisfaction. *Clinical Psychology Review*, *29*(2), 105–115. https://doi.org/10.1016/j.cpr.2008.10.004

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2022). That's a lot to process! Pitfalls of popular path models. *Advances in Methods and Practices in Psychological Science*, *5*(2). https://doi.org/10.1177/25152459221095827

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

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. https://doi.org/10.1111/j.1751-9004.2011.00355.x

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The Investment Model Scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships*, *5*(4), 357–391. https://doi.org/10.1111/j.1475-6811.1998.tb00177.x

Ryff, C., Almeida, D. M., Ayanian, J. S., Carr, D. S., Cleary, P. D., Coe, C., & Mroczek, D. K. (2007). *National survey of midlife development in the United States (MIDUS II), 2004-2006*. Inter-university Consortium for Political and Social Research [distributor].

Schmader, T., & Sedikides, C. (2018). State authenticity as fit to environment: The implications of social identity for fit, authenticity, and self-segregation. *Personality and Social Psychology Review*, *22*(3), 228–259. https://doi.org/10.1177/1088868317734080

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422

Sim, M., Kim, S.-Y., & Suh, Y. (2022). Sample size requirements for simple and complex mediation models. *Educational and Psychological Measurement*, *82*(1), 76–106. https://doi.org/10.1177/00131644211003261

Smith, L. H., & VanderWeele, T. J. (2019). Mediational E-values: Approximate sensitivity analysis for unmeasured mediator-outcome confounding. *Epidemiology*, *30*(6), 835–837. https://doi.org/10.1097/EDE.0000000000001064

Smithson, M., & Shou, Y. (2017). Moderator effects differ on alternative effect-size measures. *Behavior Research Methods*, *49*(2), 747–757. https://doi.org/10.3758/s13428-016-0735-z

Sobel, M. E. (1982) Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–321. https://doi.org/10.2307/270723

Sobel, M. E. (1996). An introduction to causal inference. *Sociological Methods & Research*, *24*(3), 353–379. https://doi.org/10.1177/0049124196024003004

Sobel, M. E. (2008). Identification of causal parameters in randomized studies with mediating variables. *Journal of Educational and Behavioral Statistics*, *33*(2), 230–251. https://doi.org/10.3102/1076998607307239

Sutton, R. I., & Staw, B. M. (1995). What theory is not. *Administrative Science Quarterly*, *40*(3), 371–384. https://www.jstor.org/stable/2393788

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. https://doi.org/10.18637/jss.v059.i05

Tofighi, D. (2021). Sensitivity analysis in nonrandomized longitudinal mediation analysis. *Frontiers in Psychology*, *12*, Article 755102. https://doi.org/10.3389/fpsyg.2021.755102

Tofighi, D., & Kelley, K. (2016). Assessing omitted confounder bias in multilevel mediation models. *Multivariate Behavioral Research*, *51*(1), 86–105. https://doi.org/10.1080/00273171.2015.1105736

Tran, U. S., Birnbaum, L., Burzler, M. A., Hegewisch, U. J. C., Ramazanova, D., & Voracek, M. (2022). Self-reported mindfulness accounts for the effects of mindfulness interventions and nonmindfulness controls on self-reported mental health: A preregistered systematic review and three-level meta-analysis of 146 randomized controlled trials. *Psychological Bulletin*, *148*(1–2), 86–106. https://doi.org/10.1037/bul0000359

VanderWeele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*, *21*(4), 540–551. https://doi.org/10.1097/EDE.0b013e3181df191c

VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford University Press.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.

Wang, Y. A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920918253

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: Questioning kappa-squared as mediation effect size measure. *Psychological Methods*, *20*(2), 193–203. https://doi.org/10.1037/met0000029

Wood, R. E., Goodman, J. S., Beckmann, N., & Cook, A. (2008). Mediation testing in management research: A review and proposals. *Organizational Research Methods*, *11*(2), 270–295. https://doi.org/10.1177/1094428106297811

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20*, 557–585. https://handle.nal.usda.gov/10113/IND43966364

Wulff, J. N., Sajons, G. B., Pogrebna, G., Lonati, S., Bastardoz, N., Banks, G. C., & Antonakis, J. (2023). Common methodological mistakes. *The Leadership Quarterly*, *34*(1), Article 101677. https://doi.org/10.1016/j.leaqua.2023.101677

Yuan, K.-H., & Maxwell, S. (2005). On the post hoc power in testing mean differences. *Journal of Educational and Behavioral Statistics*, *30*(2), 141–167. https://doi.org/10.3102/10769986030002141

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, *14*(4), 301–322. https://doi.org/10.1037/a0016972

Yzerbyt, V. Y., Muller, D., Batailler, C., & Judd, C. M. (2018). New recommendations for testing indirect effects in mediational models: The need to report and test component paths. *Journal of Personality and Social Psychology*, *115*(6), 929–943. https://doi.org/10.1037/pspa0000132

Zhang, Q., Wang, L., & Bergeman, C. S. (2018). Multilevel autoregressive mediation models: Specification, estimation, and applications. *Psychological Methods*, *23*(2), 278–297. https://doi.org/10.1037/met0000161