An Equivalence Testing Approach for Evaluating Substantial Mediation

Nataly Beribisky [a] ⓘ, Constance A. Mara [b,c] ⓘ & Robert A. Cribbie [a] ✉ ⓘ

[a]Department of Psychology, York University
[b]Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center
[c]Department of Pediatrics, University of Cincinnati College of Medicine

Abstract ■ In the past, researchers often used the nonsignificance of the direct path from the predictor to the outcome, in conjunction with a significant indirect effect, to make claims regarding 'full mediation'. However, the nil hypothesis (i.e., full mediation) is not realistic and it is well known that a nonsignificant test statistic cannot be used to establish the accuracy of a research hypothesis. In this paper, we discuss equivalence testing based procedures for assessing when a mediator explains a substantial proportion of the relationship between a predictor and an outcome. Monte Carlo simulations are used to evaluate the performance of the proposed procedure and compare it against competing alternatives, including traditional tests of full mediation and a proportion mediated approach. The proposed equivalence testing based procedures and the proportion mediated approach performed similarly across the conditions investigated. Recommendations are provided for deciding among the approaches.

Keywords ■ Mediation, equivalence testing, substantial mediation.

✉ cribbie@yorku.ca

Introduction

Mediating variables are used throughout the behavioral and social sciences to help explain and better understand the relationship between an outcome and a predictor. Mediating variables are intervening variables that transmit the effects of a predictor (X) to an outcome variable (Y). A mediating variable (M) may account for the majority of the relationship between X and Y, or may only partially account for the relationship between X and Y. Typically, a path diagram is used to depict mediation, as demonstrated in Figure 1. The indirect effect is represented by the product of paths b_{MX} (regression slope obtained from regressing M on X, often referred to as path a) and $b_{YM.X}$ (regression slope for M when regressing Y on both M and X, often referred to as path b), and the total effect of X on Y is represented by path b_{YX} (often referred to as path c). Further, the direct effect of X on Y, controlling for M, is represented by path $b_{YX.M}$ (often referred to as path c').

The number of studies investigating mediation hypotheses has increased steadily over the past couple decades, with many 'how to' papers appearing in jour-nals from many different substantive areas (e.g. Baron & Kenny, 1986; Cappelleri & Bushmakin, 2014; Caro, 2015; Frazier, Tix, & Barron, 2004; Holmbeck, 2002; Judd & Kenny, 1981; Keele, Tingley, & Yamamoto, 2015; Lachowicz, Preacher, & Kelley, 2018; Namazi & Namazi, 2016; Preacher & Hayes, 2008). Mediating variables are important because once a relationship between two variables has been established, researchers often consider the role of a third variable in this relationship (Lazarsfeld, 1955; MacKinnon, Krull, & Lockwood, 2000). There is a wide range of studies assessing whether a mediator can explain the causal relationship among two variables. For example, Ritt-Olson et al. (2005) examined whether peer influence (mediator variable) was responsible for the relationship between smoking and depression in adolescents. The researchers concluded that the relationship between smoking and depression was fully mediated by peer approval. In another study, Hadlandsmyth and Vowles (2009) examined whether the relationship between fatigue and psychosocial disability was mediated by depression. The authors concluded that depression fully mediated the relationship between fatigue and psychosocial disability. Sella,

Figure 1 ■ Baron and Kenny's (1986) representation of a three-variable mediation model. In Model A, b_{YX} represents the total effect of Y on X. In Model B, b_{MX} represents the effect of X on M, $b_{YM.X}$ represents the effect of M on Y, controlling for X, and $b_{YX.M}$ represents the direct effect of X on Y controlling for M.
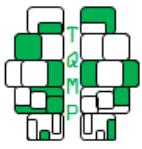
Model A

Model B
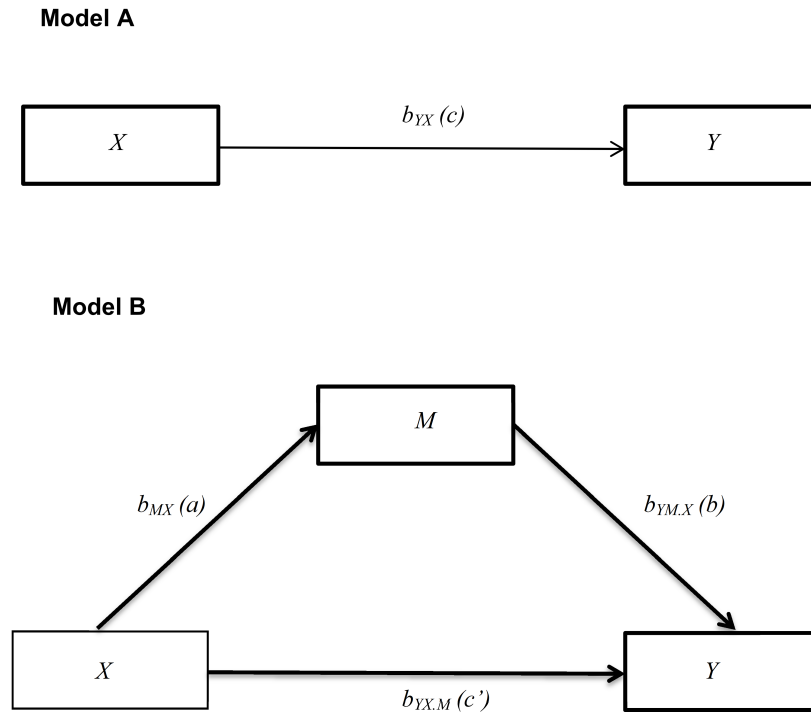
Sader, Lolliot, and Cohen Kadosh (2016) evaluated whether the relationship between basic numerical skills and mathematical expertise is fully mediated by advanced numerical skills. As a final example, Kukihara et al. (2020) explored whether resilience fully mediated the relationship between exercise/mindfulness group and psychiatric disorders. In each of these examples, the author evaluated whether the mediator fully accounted for the relationship between the predictor and the outcome.

Full (sometimes called 'complete' or 'perfect') mediation occurs when the relationship between X and Y is completely accounted for by M. In other words, M is entirely responsible for the relationship between X and Y. Accordingly, in full mediation, the relationship between X and Y disappears completely when M is controlled. In contrast, partial mediation occurs when the relationship between X and Y is reduced, but is still present when M is controlled. In other words, the relationship between X and Y is only partly accounted for by M.

For good reason, the concept of full mediation has been challenged in recent years with researchers recommending that the focus be on the strength of the indirect ef-

fect rather than on whether a variable fully or partially explains the relationship between X and Y (see Hayes, 2009; Rucker, Preacher, Tormala, & Petty, 2011). Specifically, since full/complete mediation is almost always an unrealistic goal of researchers assessing mediation, and that it is unlikely that there is zero mediation in these cases, it is assumed that almost all assessments of mediation are partial mediation cases to some degree. Accordingly, this renders the full and partial distinction ineffective.

Researchers who are interested in testing whether a mediator explains the majority of the relationship between a predictor and an outcome (rather than explaining all of the variability in the relationship between a predictor and outcome) require a procedure for evaluating this hypothesis. Therefore, the purpose of this study is to discuss a test of mediation that will evaluate whether a particular mediator accounts for a substantial proportion of the relationship between X and Y. In other words, the procedure outlined in this paper distinguishes between minimal/moderate levels of mediation and the detection of substantial mediation (as opposed to the less useful distinction between full and partial mediation). What constitutes substantial

mediation will be described below.

Our goal is to provide researchers in the behavioral and social sciences with a method that can be used to evaluate whether a third variable explains a substantial portion of the shared variability between a predictor and outcome. First, we will review traditional methods for assessing mediation. Second, we will introduce the concept of equivalence testing (the theoretical foundation on which the test of substantial mediation is built). Third, we will describe the proposed test of substantial mediation. Fourth, we will describe a (fictitious) illustrative example to guide researchers through the process of utilizing the novel mediation test (compared to traditional approaches). Fifth, we present a simulation study that evaluates the statistical properties of each of the methods described in the paper.

Traditional Methods for Assessing Full Mediation

Baron and Kenny (BK) approach. Statistical methods for analyzing mediation are widely discussed in psychology. One of the most popular methods for analyzing mediation was proposed by Judd and Kenny (1981) and Baron and Kenny (1986). Baron and Kenny (1986) suggested that assessment of mediation should be done in a series of steps: 1) Demonstrate that Y can be predicted by X (i.e., estimate and test path $b_{YX/c}$). This step establishes that there is an effect to be mediated; 2) Demonstrate that M can be predicted by X (i.e., estimate and test path $b_{MX/a}$); 3) Demonstrate that Y can be predicted from M, when both X and M are predictors in the same regression model (i.e., estimate and test path $b_{YM.X/b}$); and 4) To establish that M fully mediates the X-Y relationship, the effect of X on Y, controlling for M, should be zero (i.e., path $b_{YX.M/c'}$ should be zero), whereas to establish that M partially mediates the X-Y relationship, the effect of X on Y should be reduced (relative to the simple regression of Y on X). It is important to highlight here that full mediation is often concluded if steps 1 through 3 are satisfied, and if the test of H_0: $b^*_{YX.M} = 0$ (b^* represents the population slope) is not rejected (Rucker et al., 2011). However, it is well known that nonrejection of a null hypothesis cannot be used to support the null as the true effect (as will be explained in detail when we introduce equivalence testing).

Structural equation modeling (SEM) approach. James and Brett (1984) recommend the use of structural equation modeling for testing hypotheses of mediation. Unlike regression approaches, SEM uses a maximum likelihood estimator (compared to least squares in regression) and tests a wider range of models. The differences between the two approaches have been contended (e.g. Hayes, Montoya, & Rockwood, 2017; Pek & Hoyle, 2016). James and Brett (1984), and James, Mulaik, and Brett (2006) argue for the use of a full mediation model as the underlying model
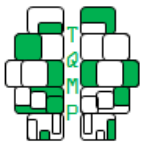
for assessing mediation (see Figure 2), as there is no requirement that X and Y are related a priori. Thus, this model does not test for the existence of a direct effect of X on Y, and instead focuses on the indirect effect being tested. Mediation is established if both the paths from X to M, and from M to Y, are nonzero, and if the model fits the data using a goodness of fit test (e.g., likelihood ratio test). As there is only one degree of freedom in the full mediation model, the goodness of fit test assesses whether the direct effect from X to Y is zero. As with the BK method, it is important to note that full mediation is typically said to exist if the paths from X to M, and from M to Y, are statistically significant and if the test of H_0: $b^*_{YX.M} = 0$ is not rejected (which in this case results from a nonsignificant likelihood ratio χ^2 test). If a researcher is instead interested in testing a partial mediation model, a direct effect from X to Y can be included to evaluate whether M partially mediates the relationship between X and Y. For partial mediation to be established, all paths in the model (Y on X, Y on M, and M on X) should be nonzero.

Iacobucci, Saldanha, and Deng (2007) conducted simulation studies comparing the BK and SEM methods for testing mediation. They found that across all sample sizes (including sample sizes as small as 30) and across different degrees of mediation, that the SEM approach was slightly superior to the Baron and Kenny regression method for correctly detecting mediation. These authors argue that the SEM approach should always be employed when testing hypotheses of mediation because of the parsimony of fitting a single model which estimates all the parameters in the model simultaneously.

When assessing mediation via BK or SEM, researchers sometimes supplement the results of a BK or SEM approach with a statistical test of the significance of the indirect effect of X on Y through M (i.e., test the significance of $b_{MX}b_{YM.X}$). Current recommendations suggest that the test of the indirect effect should be the only test conducted, not merely a supplement (see Hayes, 2009; Hayes et al., 2017). This test was traditionally done using a test proposed by Sobel (1982, 1986), although more recently, the statistical significance of the indirect effect is assessed using more robust techniques, such as percentile bootstrapped confidence intervals (Fritz, Taylor, & MacKinnon, 2012; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher & Hayes, 2004).

Problems with Assessing "Full" Mediation

One important limitation of the way in which full mediation is assessed with the BK and SEM approaches is the use of a nonsignificant test statistic as the primary criteria for declaring full mediation. In the BK approach, a lack of statistical significance of X as a predictor of Y, control-

Figure 2 ∎ SEM representation of a three-variable mediation model, which uses the full mediation model as the default model. In this model, b_{MX} represents the effect of X on M, and b_{YM} represents the effect of M on Y.
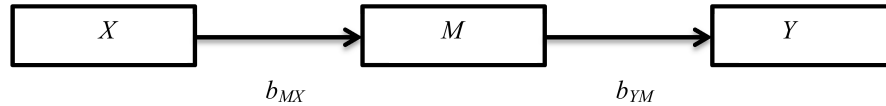
ling for M (after finding statistically significant relationships in the first three steps of the BK approach), is often used to establish full mediation. With the SEM approach, a lack of statistical significance of a goodness of fit statistic (e.g., the maximum likelihood based chi-square test, which is the default goodness of fit test for most SEM software) is often used to establish full mediation (this is theoretically, but not statistically, the same test that is conducted with the BK approach, namely a test of the lack of statistical significance of the direct relationship from X to Y, when M is also a predictor of Y). There are (at least) two problems with this approach: 1) not rejecting the null hypothesis of a lack of a direct effect from X to Y does not establish that the relationship between X and Y is zero; and 2) power for detecting full mediation (i.e., not rejecting the null hypothesis of a lack of a direct effect from X to Y) is backwards; in other words, power is increased by reducing, rather than increasing, sample size (Rucker et al., 2011).

Shrout and Bolger (2002) also discuss the limitations of establishing full mediation via null hypothesis testing, and subsequently discuss the use of confidence intervals and an estimate of the proportion of the total variability that is mediated as alternatives to using the lack of statistical significance of the direct effect as evidence of mediation. Although this approach is beneficial in that it attempts to move the discussion away from the full versus partial mediation distinction when assessing the potential importance of a single mediator, this does not provide psychological researchers with a valid test of whether a mediator explains a substantial proportion of the variability in the relationship between a predictor and an outcome. Another alternative proposed by Kenny (2020) asserts that full mediation can be concluded if $\frac{b_{mx}b_{ym.x}}{b_{yx}} > .80$ and $|\beta_{yx}| > .2$, where β represents the standardized coefficient (i.e., the regression coefficient if all variables are standardized a priori). The advantage of this approach is that it is not based on traditional null hypothesis testing, and, since the cutoff for the proportion mediated is less than 1, it is not a test of full mediation (and is actually more related to the test of substantial mediation proposed in this paper). However, to the best of our knowledge, this approach has not been evaluated. Also see Preacher and Kelley (2011) for a

discussion of the issues with estimating the proportion mediated. This approach is referred to below as the K method.
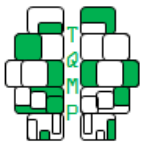
Novel Approach for Assessing Substantial Mediation

In order to motivate a novel approach to assessing mediation, it is helpful to look at the BK and SEM approaches to assessing full mediation from a slightly different perspective. A test of whether the direct effect from X to Y is zero is comparable to asking whether the regression coefficient for predicting Y from X (i.e., b_{YX}) is equal to a coefficient reproduced by the indirect effect ($b_{MX}b_{YM.X}$). If $b_{YX} = b_{MX}b_{YM.X}$ then all of the XY relationship is mediated (i.e., the indirect effect of X on Y through M explains the observed relationship between X and Y). This hypothesis could be tested using a statistical test of the difference between regression coefficients (e.g. Steiger, 1980; MacKinnon et al., 2002), which is an alternative method for assessing full mediation that has been outlined in James et al. (2006). However, this approach shares the same disadvantage of the BK and SEM approaches to assessing mediation discussed above, namely that full mediation would be demonstrated if the test of the difference between regression coefficients was not statistically significant.

A novel test of mediation can be established by finding an appropriate way to determine the 'equivalence' of b_{YX} and $b_{MX}b_{YM.X}$ or, equivalently, that $b_{YX.M}$ is negligible. By incorporating equivalence testing, a researcher could move away from looking at full mediation and instead focus on whether a particular mediator accounts for the majority of the variability in the relationship between X and Y (i.e., substantial mediation).

Equivalence Testing

The backbone to this proposed approach is the framework of equivalence testing, which statistically tests for a negligible relationship among variables. This approach was popularized in biopharmaceutical research where the goal was to demonstrate the bioequivalence of drugs, but was introduced to the behavioral and social sciences literature through influential articles by Rogers, Howard, and Vessey (1993), Seaman and Serlin (1998), Tryon (2001), and others. Suppose that a researcher is interested in demonstrating

that two population means are equivalent on an outcome. Unlike traditional difference-based tests, the goal of equivalence testing is to show that the difference between the means of the two populations is too small to be considered meaningful. In other words, researchers do not need to show that the means are 'identical' (as with the traditional point null hypothesis, H_0: $\mu_1 = \mu_2$), but only that any differences between the means are inconsequential.

Several tests have been designed to evaluate the equivalence of two population means, with Schuirmann's (1987) two-one-sided tests (TOST) of equivalence being one of the most popular. The first step with Schuirmann's TOST of equivalence is to establish a critical mean difference (or equivalence interval) for declaring two population parameters (e.g., means, regression parameters) equivalent ($-\delta$, δ). Any mean difference that falls within the equivalence interval would be considered meaningless within the context of the study. Two simultaneous one-sided hypothesis tests can be used to establish equivalence, where the null hypothesis relates to the non-equivalence of the population means and can be expressed as two separate composite hypotheses. This is equivalent to determining if the $(1-2\alpha)$% confidence interval for the parameter is contained completely within the user-specified equivalence interval. For example, if we are evaluating the equivalence of two population means, we can represent the hypotheses as H_{01}: $\mu_1 - \mu_2 \geq \delta$ and H_{02}: $\mu1 - \mu_2 \leq -\delta$. Rejection of H_{01} implies that $\mu_1 - \mu_2 < \delta$, and rejection of H_{02} implies that $\mu_1 - \mu_2 > -\delta$. Further, rejection of both hypotheses implies that the difference in the means falls within the bounds from $-\delta$ to δ and the means can be considered equivalent.

The choice of an equivalence interval will vary greatly based upon the nature of the study. Although both Rogers et al. (1993) and Cohen (1988) provide possibilities for δ (e.g., 20% difference between two means, lower bound of the cutoff for a small standardized effect size, such as a $d = 0.2$ or $r = 0.1$), they discouraged researchers from using them as strict guidelines as the specification of δ depends heavily upon the substantive research question.

Equivalence Test of Substantial Mediation (ESM)

Earlier papers on equivalence testing focused on establishing the equivalence of means (e.g. Cribbie, Gruman, & Arpin-Cribbie, 2004; Rogers et al., 1993; Seaman & Serlin, 1998; Tryon, 2001), although more recent research has extended the approach to assessing a lack of association among variables (e.g. Goertzen & Cribbie, 2010), a lack of interaction (Cribbie, Ragoonanan, & Counsell, 2016), etc. (see Wellek, 2010, for an extensive coverage of equivalence testing). Borrowing from the logic of lack of association tests, we propose an equivalence test for the direct effect

to assess substantial mediation. Substantial mediation testing is an equivalence-testing based procedure that evaluates whether a variable is able to explain a considerable (i.e., substantial) proportion of the variability in the relationship between a predictor and an outcome. It is important to point out that this is distinct from assessing the statistical significance of the indirect effect, where the focus is only on the indirect effect (and no interest is paid to the distribution of variability across the indirect and direct effects). With the proposed method, we can estimate the equivalence of b_{YX} and $b_{MX}b_{YM.X}$ from the direct effect of X upon Y, $b_{YX.M}$ (Preacher & Kelley, 2011). As discussed above, this can be evaluated by demonstrating that the $(1-2\alpha)$% confidence interval for $b_{YX.M}$ falls within the equivalence interval from $-\delta$ to δ, where δ represents a pre-specified difference between the coefficients b_{YX} and $b_{MX}b_{YM.X}$ (or equivalently the magnitude of $b_{YX.M}$) that would be considered meaningless within the framework of the study.

What magnitude of the direct effect is considered 'meaningless' depends highly on the nature of the study and could be framed in different metrics. Like an interpretation of an effect size, this choice of an effect will depend on the context of the study, the measures used, and what is pertinent to the research question. For example, one researcher may decide that a raw direct effect of less than .1 would be considered inconsequential, whereas another researcher may find that 20% of the standard error might be appropriate. Another option is to standardize the variables and then frame the question in terms of what standardized regression coefficient is meaningful. What is important to note is that increasing the size of δ will increase power. However that power increase comes at the cost of declaring equivalence via a larger interval. Another way to put this is that with a large enough interval any confidence interval around the direct effect could be declared equivalent; however, there would be little confidence that the effect is actually negligible.

Following Schuirmann's (1987) TOST approach, the null and alternate hypotheses for this test are H_{01}: $b^*_{YX.M} \geq \delta$ | H_{11}: $b^*_{YX.M} < \delta$ and H_{02}: $b^*_{YX.M} \leq -\delta$ | H_{12}: $b^*_{YX.M} > -\delta$. Rejection of H_{01} implies that the direct effect is less than δ and rejection of H_{02} implies that the direct effect is greater than $-\delta$. Rejection of both null hypotheses implies that the direct effect falls within the interval $(-\delta, \delta)$. Rejection of both hypotheses is required in order to establish equivalence, and infer that the mediator explains a substantial proportion of the variability in the outcome. Again, this is equivalent to demonstrating that the $100(1-2\alpha)$% confidence interval for the direct effect falls completely within the equivalence interval. An important requirement of the procedure is that $|\beta_{YX}| > 0$ or
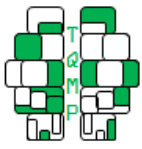
$|\beta_{MX}\beta_{YM.X}| > 0$; in other words, if the total effect and indirect effect are near zero, then it would not make sense to conclude that the indirect effect accounts for a substantial proportion of the total effect. Although what magnitude of effect is required will depend on the context of the study, we suggest that $|\beta_{YX}| > .15$ or $|\beta_{MX}\beta_{YM.X}| > .15$ as possibilities, and thus we propose two versions of the procedure. For the ESM_{tot}, substantial mediation is concluded if $\beta_{YX} > .15$ and the $1 - 2\alpha$ percentile bootstrap confidence interval for $\beta_{YX.M}$ falls completely within the equivalence interval, and for the ESM_{ind}, substantial mediation is concluded if $\beta_{MX}\beta_{YM.X} > .15$ and the $1 - 2\alpha$ percentile bootstrap confidence interval for $\beta_{YX.M}$ falls completely within the equivalence interval.

Illustrative Example

To illustrate the use of the ESM and alternative methods, we use simulated data based on the research hypothesis of Hadlandsmyth and Vowles (2009). Specifically, our interest was in whether the relationship between fatigue (X) and psychosocial disability (Y) was substantially mediated by depression (M). The data, $N = 900$, very closely matched the correlation structure of the original data. The code for the illustrative example is available at https://osf.io/75a3x/.

Before demonstrating the procedures described above, it is informative to also look at the effect sizes associated with this effect. Thus, before we present detailed results of the different mediation approaches, we use upsilon ($\hat{v}$) and adjusted upsilon ($\tilde{v}$) as effect size measures which are equivalent to the squared completely standardized indirect effect (Lachowicz et al., 2018). We also use the proportion of the total effect which is mediated ($\frac{b_{mx}b_{ym.x}}{b_{yx}}$) which describes the proportion of the total effect of fatigue on psychosocial disability which is mediated by depression. All of these measures are available in the MBESS package in R (Kelley, 2007). Accordingly, the squared completely standardized indirect effect is $\hat{v}$ = 6.57%, 95% CI [4.66%, 8.87%] or $\tilde{v}$ = 6.52%, 95% CI [4.61%, 8.82%]. The proportion of the total effect mediated is 66.42%.

For the BK approach, we followed the traditional procedure outlined above that involves running three regression models, the first regressing Y on X, the second regressing M on X, and the third regressing Y on M and X. For the SEM approach we utilized the model in Figure 2, regressing Y on M, and M on X.

For the ESM and K approaches we used the R function esm (available at https://osf.io/75a3x/), where we specified the standardized solution and a smallest meaningful standardized coefficient of .2; in other words, for the ESM procedures, we would declare that substantial mediation has occurred if the standardized coefficient for the direct effect of fatigue on psychosocial disability falls within the in-

terval $\{-.20, .20\}$ and the other conditions are met. This would indicate that a substantial proportion of the relationship between fatigue and psychosocial disability can be explained by depression. For the ESM_{tot}, substantial mediation is concluded if $|\beta_{YX}| > .15$ and the $1 - 2\alpha$ percentile bootstrap confidence interval (in this case $\alpha = .05$, so the 90% confidence interval) for $\beta_{YX.M}$ falls completely within the equivalence interval $\{-.20, .20\}$.
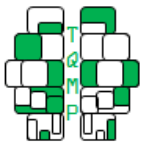
In this example, $\beta_{YX} = .386$ ($p < .001$), $\beta_{MX} = .409$ ($p < .001$), $\beta_{YM.X} = .627$ ($p < .001$), $\beta_{YX.M} = .129$ ($p < .001$), and the 90% percentile bootstrap confidence interval for $\beta_{YX.M}$ is {.090, .179}. Thus, following the steps for the BK, we cannot conclude full mediation because even though the β_{YX}, β_{MX}, and $\beta_{YM.X}$ paths are significant, the $\beta_{YX.M}$ path is also significant. The same conclusion is found for the SEM approach, where although the β_{YM} and β_{MX} paths are significant, the χ^2 statistic is also significant [$\chi^2(1 \text{ df}) = 23.735$, p < .001] indicating that $\beta_{YX.M}$ is nonnull.

For the ESM approach, $|\beta_{YX}| > .15$ and the 90% confidence interval for $\beta_{YX.M}$ falls completely within the equivalence interval, and thus the necessary conditions are met and we can conclude that depression substantially mediates the relationship between fatigue and psychosocial disability.

For the K procedure, we can conclude that a substantial portion of the variability is due to mediation if $\frac{b_{mx}b_{ym.x}}{b_{yx}}$ > .80 and $|\beta_{YX}| > .2$. In this example, $\frac{b_{MX}b_{YM.X}}{b_{YX}} = \frac{\beta_{MX}\beta_{YM.X}}{\beta_{YX}} = \frac{(.409)(.627)}{.386} = .664$ and, as presented above, $\beta_{YX} = .386$. Thus, we cannot conclude that depression completely (or fully) mediates the relationship between fatigue and psychosocial disability since, although $\beta_{YX} > .2$, $\frac{b_{mx}b_{ym.x}}{b_{yx}} < .8$.

Monte Carlo Study

A Monte Carlo study was used to compare the ESM procedure proposed in this study with the previously proposed BK and SEM methods for detecting full mediation and the proportion of variability (K) method. A Monte Carlo study simulates data that are representative of the nature of the data found in the discipline, and then evaluates the performance of the relevant procedures. In this study, we simulate data with varying degrees of mediation and then compare the performance of the relevant procedures across these conditions (and other conditions described below). It is important to point out that the methods are testing different research hypotheses; the BK and SEM approaches are testing whether the direct effect of X on Y is not statistically significant, the K approach is testing whether the proportion of variability attributed to the indirect effect is large, and the ESM approach is testing whether the di-

rect effect of X on Y falls within the equivalence interval. For the ESM approach, both the ESM_{tot} and ESM_{ind} were evaluated. As discussed above, both approaches protect the researcher from concluding substantial mediation when in fact there is no total or indirect effect, respectively. In both cases, the standard errors were derived using the percentile bootstrap approach (500 bootstraps).

The open-source statistical package R (R Development Core Team, 2020) was used to conduct the simulations and analyses. The approach proposed by Caron and Valois (2018) for simulating three variable mediation models was adopted. The `lavaan` package in R was employed to perform the SEM analysis on the mediation model (Rosseel, 2012). Following Caron and Valois (2018), the following equations were used to generate standardized X, M, and Y, using the `rnorm` function in R for random normal variables (note that since X, M and Y are standardized we could equivalently replace b with β). We start by simulating the mediator (M) as:

$$M_i = b^*_{MX} X_i + e_{1i}, \tag{1}$$

where X_i is a normally distributed variable with $\mu = 0$ and $\sigma = 1$, and e_{1i} is a normally distributed variable with $\mu = 0$ and $\sigma = \sqrt{1 - (b^*_{MX})^2}$. Y_i is simulated as:

$$Y_i = b^*_{YX.M} X_i + b^*_{YM.X} M_i + e_{2i}, \tag{2}$$

where e_{2i} is a normally distributed variable with $\mu = 0$ and

$$\sigma = \sqrt{1 - \left[(b^*_{YX.M})^2 + (b^*_{YM.X})^2 - 2b^*_{MX} b^*_{YX.M} b^*_{YM.X} \right]}. \tag{3}$$

The equivalence interval $(-\delta, \delta)$ for the ESM was set to $(-.15, .15)$. As noted previously, the establishment of an appropriate equivalence interval depends on the nature of the study, and thus, in the absence of a substantive problem, this interval was selected somewhat arbitrarily. Increasing or decreasing the size of δ would result in a respective increase or decrease, respectively, in the power of the ESM test and therefore we did not explore multiple conditions.

Two primary factors were manipulated in this study, sample size and effect size. Sample sizes were set at 50, 100, 200, 350, 500 and 1500, chosen to be comparable to the sample sizes common within the social and behavioral sciences. Regression coefficients for the b^*_{MX} and $b^*_{YM.X}$ paths from Figures 1 and 2 were each set at 0 ($b^*_{MX} b^*_{YM.X} = 0$), .30 ($b^*_{MX} b^*_{YM.X} = .09$), .50 ($b^*_{MX} b^*_{YM.X} = .25$), .60 ($b^*_{MX} b^*_{YM.X} = .36$) or .70 ($b^*_{MX} b^*_{YM.X} = .49$). In order to also investigate the effect of different b^*_{MX} and $b^*_{YM.X}$ coefficients on the results, we also investigated conditions in which one coefficient was larger than the other. Specifically, we looked

at cases where one coefficient was 0 and one coefficient was .50 (for $b^*_{MX} b^*_{YM.X} = 0$), where one coefficient was .173 and one coefficient was .519 (for $b^*_{MX} b^*_{YM.X} = .09$), where one coefficient was .353 and one coefficient was .709 (for $b^*_{MX} b^*_{YM.X} = .25$), where one coefficient was .490 and one coefficient was .735 (for $b^*_{MX} b^*_{YM.X} = .36$), and where one coefficient was .606 and one coefficient was .808 (for $b^*_{MX} b^*_{YM.X} = .49$). Note that the ratio of these coefficients (0/.50 = 0, .173/.519 $\approx$ .33, .353/.709 $\approx$ .5, .490/.735 $\approx$ .66, .606/.808 $\approx$ .75) is proportional to the distance of the coefficients to the ceiling (i.e., greater variability in coefficients is permissible when the values do not approach a ceiling effect).

Regression coefficients for the $b^*_{YX.M}$ path were set at 0, .05, or .25. For $b^*_{MX} b^*_{YM.X} = .25$, .36, and .49, when $b^*_{YX.M} = 0$, each of full mediation (BK/SEM), substantial mediation (ESM_{tot}/ ESM_{ind}), and proportion mediated > .80 are present and thus the simulation study is used to evaluate the power of each of the procedures (although it is important to recall that the K procedure is not an inferential procedure and therefore does not have strictly defined Type I and/or power rates). For $b^*_{MX} b^*_{YM.X} = .25$, .36, and .49, when $b^*_{YX.M} = .05$, this is substantial mediation because the direct effect falls within the equivalence interval (-.15, .15), so this is still a power condition for the ESM_{tot}/ ESM_{ind} procedures. However, the BK and SEM approaches should reject the null hypothesis, H_0: $b^*_{YX.M} = 0$, and indicate that full mediation is not occurring. For the K procedure, the proportion mediated exceeds .80 and thus these are also power conditions. For $b^*_{MX} b^*_{YM.X} = .25$, .36, and .49, $b^*_{YX.M} = .25$, again, the BK and SEM approaches should reject the null hypothesis H_0: $b^*_{YX.M} = 0$ and indicate that full mediation is not occurring. The same general situation exists for the K procedure (proportion mediated is not > .80 in any condition so full mediation should not be concluded), however again recall that there is no formal hypothesis testing for the K procedure. For the ESM_{tot}/ ESM_{ind} procedures, the direct effect falls outside the equivalence interval (-.15, .15) when $b^*_{YX.M} = .25$, and therefore this test should not indicate that M is a substantial mediator (i.e., rejecting the null hypothesis is a Type I error).

It is worth noting that when $b^*_{MX} b^*_{YM.X} = 0$, there is no indirect effect, thus, this condition is analogous to testing error rates for the different procedures as the indirect effect itself is not responsible for the effect of $b^*_{YX.M}$. An anonymous reviewer suggested the $b^*_{MX} b^*_{YM.X} = .09$ condition and this is a complicated one. Although the same conclusions regarding substantial mediation exist as for $b^*_{MX} b^*_{YM.X} = 0$, .25, .36, and .49, because the total effect (.14) or indirect effect (.09) are not greater than the cutoff (.15), testing for substantial mediation is not conducted in
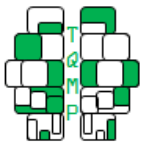
most cases (although with sampling variance in some cases the total/indirect effect might exceed .15).

The nominal Type I error rate (α) was set to .05 for all conditions. For evaluating the power of the procedures, we conducted 1000 simulations for each condition.

Results

Power rates for the BK and SEM approaches were almost identical (except for conditions with small N and small $b^*_{YX.M}$) and therefore the results will be discussed together. Similarly, unless otherwise specified, the rejection rates for the two ESM procedures (ESM_{tot}/ESM_{ind}) were almost identical across all conditions and therefore will be discussed together under 'ESM'. Power rates for the BK, SEM, K and both ESM procedures for $b^*_{MX}b^*_{YM.X}$ = 0, .09, .25, .36 and .49 are presented in Tables 1, 2, 3, 4, and 5, respectively, at the end of this article.

Validity Check for the ESM Procedure

To evaluate whether the ESM procedure provided a valid test of the composite null hypothesis (H_{01} and H_{02}), the empirical Type I error rate for the procedure was compared with the nominal alpha level (.05) using a sample size of $N = 500$ and $b^*_{MX} = b^*_{YM.X} = .5$. To evaluate the empirical Type I error rate, it was necessary to set $b^*_{YX.M}$ ($\beta^*_{YX.M}$) equal to δ. The empirical Type I error rate was .054, which deviated only slightly from the nominal level.

Full/Substantial Mediation Present ($\beta_{YX.M} = 0$)

Given $b^*_{MX}b^*_{YM.X}$ =.25, .36, and .49, it is important to note that this is a power condition for the ESM because $\beta^*_{YX.M}$ is less than δ. The BK/SEM approaches had greater power than the ESM approaches, until the sample size was very large. For example, the power rates for the ESM procedures were close to 0 until the sample size was greater than $N = 100$, and approached unity as the sample size reached $N = 1500$. On the other hand, the power rates for the BK/SEM increased from $N = 50$ to $N = 1500$, but the results at lower sample sizes were often much greater than 0. It is important to note that the power rates for the BK/SEM approaches reach a ceiling of $1 - \alpha$ because the test is measuring the probability of not declaring $b_{YX.M}$ different from 0 (or, as discussed earlier, not rejecting H_0: $b^*_{YX.M}$ = 0), which, with a large enough sample size, is simply $1 - \alpha$. When $b^*_{MX}b^*_{YM.X}$ = .25, the K procedure had power rates that were between the ESM and BK/SEM procedures when sample size was between $N = 50$ and $N = 350$, but, like the other procedures, approached unity at $N = 500$ and $N = 1500$. When $b^*_{MX}b^*_{YM.X}$ = .36/.49 the rates increased faster with N, with BK/SEM and the K methods approaching .9 and the ESM methods approaching .7 at $N = 100$. When $b^*_{MX} > b^*_{YX.M}$ the power rates

are generally lower than when $b^*_{MX} < b^*_{YX.M}$ for all procedures excluding SEM. This effect was more pronounced for the ESM procedures.

For the small indirect effect of $b^*_{MX}b^*_{YM.X}$ = .09, the ESM approaches did not attain high power, even at larger sample sizes (recall that the minimum requirement for the total/indirect effects, .15, was not met and thus in most conditions substantial mediation testing was not conducted). The ESM_{tot} approach had slightly higher power than the ESM_{ind} approach. The K approach had low power at all sample sizes, but rates decreased to 0 as the sample size approached 1500. Like the ESM procedures, the minimum β_{yx} relationship requirement was not met. The BK/SEM approaches had low power at $N = 50$ (.014 to .36) but approached .95 by $N = 1500$.

The ESM procedures almost never declared full mediation when $b^*_{MX}b^*_{YM.X}$ = 0 (the only non-zero rate was 0.10% in the $N = 200$, $b^*_{MX} < b^*_{YX.M}$ condition). The K procedure had similar rates to those of the ESM procedures. In contrast, the BK/SEM procedures declared full mediation when $b^*_{MX}b^*_{YM.X}$ = 0 up to 3.10% of the time.

Substantial Mediation Present ($b^*_{YX.M}$ = .05)

When $b^*_{YX.M}$ = .05, again the direct effect is within the equivalence interval and thus this is a power condition for the ESM. Given that the BK approach is assessing whether the test of H_0: $b^*_{YX.M}$= 0 is not rejected, and the SEM approach, because there is only one degree of freedom in the full mediation model, similarly assesses whether the direct effect from X to Y is zero via the goodness of fit test, this condition should fail these assessments and not indicate that full mediation is present (i.e., any nonrejections are Type II errors). Similar to the previous condition, for $b^*_{MX}b^*_{YM.X}$ =.25, .36, and .49, the ESM procedures have almost zero power until $N = 100$, but approach unity at $N = 1500$. Incorrect declarations of full mediation for the BK/SEM approaches increased from $N = 50$ to $N = 100/200$, then decreased as sample sizes increased towards $N = 1500$. These error rates are a function of the requirement by these approaches that there are both significant effects (b_{MX}, $b_{YM.X}$, and b_{YX}) and nonsignificant effects ($b_{YX.M}$). Thus, at the larger sample sizes, there is high power for rejecting H_0: $b^*_{YX.M}$ = 0 and also for rejecting H_0: $b^*_{YM.X}$ = 0 and H_0: b^*_{MX} = 0. At $N = 1500$, even though $b^*_{YX.M}$ is not zero, there is still approximately up to a 68% chance of declaring full mediation using the BK/SEM procedures. When $b^*_{MX} > b^*_{YX.M}$ the power rates are lower than when $b^*_{MX} < b^*_{YX.M}$ for ESM/K, whereas the reverse is true for BK/SEM at $N > 100$.

At $b^*_{MX}b^*_{YM.X}$ =.25, .36, and .49, like in the full mediation condition, for smaller sample sizes the K procedure's power rates were between BK/SEM and both ESM proce-

dures. At moderate sample sizes the power rates for K were similar to the ESM procedures at smaller $b^*_{MX}b^*_{YX.M}$ and greater than the ESM procedures at larger $b^*_{MX}b^*_{YX.M}$. Further, at small to moderate sample sizes the rates for K were lower than that of the BK/SEM procedures, whereas at large sample sizes the rates for K were larger than that of the BK/SEM procedures. This makes sense since the K procedure is based on an effect size, whereas the BK/SEM procedures are based on hypothesis testing; at large N the BK/SEM procedures have high power for rejecting H_0: $b^*_{YX.M}$ = 0 even though the effect size is very small.

When $b^*_{MX}b^*_{YM.X}$ = .09, the ESM approaches had low power across all sample sizes. As discussed earlier, the minimum requirement for the total/indirect effects (.15) was not met and thus in most conditions substantial mediation testing was not conducted. Power was notably higher for the ESM_{tot} approach compared to the ESM_{ind} procedure, since the total effect was closer to the cutoff. For instance, at $N = 1500$, power was between .279 and .356 for ESM_{tot} but was at 0 for ESM_{ind}. Like for $b^*_{YX.M}$ = 0, the K procedure had consistently low rates since neither of the conditions for the procedure (proportion mediated > .80 and a total effect > 2) were met (i.e., the procedure performed as it should). Power declarations increased for the BK/SEM procedure until $N = 500$ and then fell when $N = 1500$; this is a similar pattern to the other indirect effect conditions with power for the b_{MX}, $b_{YM.X}$, and b_{YX} effects increasing with sample size, but power for the (near zero) direct effect being lower at $N < 1500$.
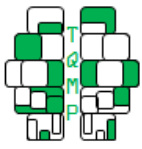
The ESM procedures (appropriately) almost never declared full mediation when $b^*_{MX}b^*_{YM.X}$ = 0 (the only non-zero rate was 0.10% in the $N = 200$, $b^*_{MX} < b^*_{YX.M}$ condition). The error rates for BK/SEM were higher, reaching 5.6%. The K procedure, also appropriately, almost never declared substantial mediation.

Generally, when $b^*_{YX.M}$ = .05, there were many instances of increased power for detecting mediation (relative to $b^*_{YX.M}$ = 0 condition). One reason for this power increase is that fixing $b^*_{YX.M}$ to .05 increases the overall relationship between X and Y (total effect). In contrast, when fixing $b^*_{YX.M}$ to 0, the total effect (β_{YX}) may still not be large enough to be statistically significant (especially at smaller samples), meet the minimum total effect requirement for the ESM_{tot} and K procedures, or indirectly inflate the minimum indirect effect requirement for the ESM_{ind} procedure.

To ensure that it was indeed the minimum total or indirect effect requirement that was inflating the power rates at $b^*_{YX.M}$ = .05 for the ESM procedures (as opposed to the confidence interval requirement), we recorded the proportion of instances where only the minimum effect requirement for either procedure was met at $N = 50$, for $b^*_{YX.M}$
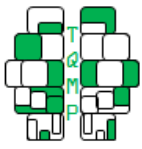
= .0 and $b^*_{YX.M}$ = .05. We found that the minimum effect requirement was met more frequently at $b^*_{YX.M}$ = .05 than $b^*_{YX.M}$ = .0. Secondly, we reran conditions in which there was no requirement of a minimum underlying total or indirect effect for the ESM procedures. We found that the power increase at $b^*_{YX.M}$ = .05 disappeared for both methods. We see less instances of this power bump for all the procedures when sample sizes increase (less cases at $N = 1500$ compared to $N = 50$) and when the magnitude of the indirect effect becomes larger (less cases with $b^*_{MX}b^*_{YM.X}$ = .49 compared to $b^*_{MX}b^*_{YM.X}$ = .09).

Full/Substantial Mediation Not Present ($b^*_{YX.M}$ = .25)

When $b^*_{YX.M} > \delta$, full or substantial mediation does not exist and therefore any indications of full mediation by the BK or SEM approaches, substantial mediation by the ESM approaches, or proportion mediated > .8 by the K approach are errors. The ESM procedures almost never indicated substantial mediation to be present, when in fact it was not present, in any condition (the largest rate observed was .004). On the other hand, the BK/SEM and, to a lesser extent, K approaches were only error free at larger sample sizes. At the smaller sample sizes, the BK/SEM approaches often found full mediation to exist when in fact it did not with error rates lower at larger values of $b^*_{MX}b^*_{YM.X}$. For example, for $b^*_{MX} = b^*_{YM.X}$ = .5 and $N = 50$ the BK/SEM approaches sometimes incorrectly indicated full mediation more than 50% of the time. To a lesser degree, the K approach in the same conditions indicated that full mediation existed up to 20% of the time.

It is important to highlight that the error rates when $b^*_{YX.M} > 0$, as with $b^*_{YX.M}$ = 0, are complicated because they are a function of the requirement by these traditional approaches that there are significant effects (b_{MX}, $b_{YM.X}$, and b_{YX}) and nonsignificant effects ($b_{YX.M}$). At the smaller sample sizes, H_0: $b^*_{YX.M}$= 0 is rarely rejected (which is one of the necessary requirements for full mediation with traditional approaches) because there is very little statistical power for testing this relationship. However, there is also reduced statistical power for rejecting H_0: $b^*_{YM.X}$ = 0 and H_0: b^*_{MX} = 0 or assessing the proportion of variability. Therefore, it is difficult to make sense of the pattern of error rates for the BK/SEM approaches until sample sizes become large and the power for rejecting H_0: $b^*_{YX.M}$= 0 approaches unity; however, in general for this condition the consequence is that many false declarations of full mediation are more likely with the BK/SEM methods at smaller sample sizes.

The ESM procedures correctly never declared full mediation when $b^*_{MX}b^*_{YM.X}$ = 0. Error rates for the BK/SEM procedure were low at large sample sizes ($N > 300$), but went as high as 9.50% when $N = 50$. The K approach ap-

propriately never declared full mediation in this condition.

Discussion

Researchers in psychology are commonly interested in exploring potential mediators of proposed relationships. As outlined in the introduction, these explorations can involve an assessment of whether a mediator fully or completely determines the relationship between a predictor and an outcome variable. Historically, researchers used a lack of statistical significance of the direct effect between the predictor and outcome, after controlling for the mediator, as evidence of full mediation. Although the language of 'full' and 'partial' mediation has been found to be of little use in understanding the statistical significance of the indirect effect, researchers are still concerned with assessing these forms of mediation in order to understand the nature of the mediator they are interested in. For example, studies such as those by Ritt-Olson et al. (2005), Hadlandsmyth and Vowles (2009), Sella et al. (2016), and Kukihara et al. (2020), are theoretically interested in whether a single variable can account for the majority of the relationship between a predictor and an outcome, and this information is best understood in terms of whether the mediating variable can substantively account for the relationship between an independent variable and an outcome.
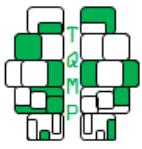
This study proposed two novel tests of substantial mediation based on equivalence testing and compared it to traditionally used methods for assessing full mediation. The ESM methods require the researcher to specify a priori what magnitude for the direct effect would be considered inconsequential within the nature of the study. Although this a priori determination can be challenging, researchers regularly consider what level of effect size is meaningful within the context of their research question and previous literature (Beribisky, Davidson, & Cribbie, 2019). The selection of an equivalence interval is closely related to this process - a researcher must consider, within the context of their research question and the extant literature, what effect represents an appropriate inconsequential difference.

Although the proposed ESM tests have reduced power for detecting substantial mediation at smaller sample sizes, they also do not have the disadvantages shared by the BK and SEM approaches of violating the principles of null hypothesis testing (i.e., relying on nonrejection of the null hypothesis) and commonly declaring full mediation when it does not exist. More specifically, power for detecting substantial mediation with the ESM approaches (for the conditions investigated in this paper) was low at small sample sizes, and only reached acceptable levels at $N = 350$ for indirect effects of .25, .36 and .49, which is definitely a limiting factor for many researchers (although note that a fairly strict equivalence interval was used given the context-free

nature of the simulation study). The best power rates for the ESM procedures occurred when $b_{YM.X} > b_{MX}$ due to a narrower confidence interval around the direct effect when the mediator was more closely related to the outcome variable than the independent variable. This result is related to collinearity and a decreased standard error when the mediator is more closely associated to the outcome (see Kenny & Judd, 2013).

On the other hand, with $N = 50$, the probability of declaring "full" mediation when it was not present reached as high as 94% with the BK and SEM procedures, and as high as 93% with $N = 100$ when $\beta_{MX} > \beta_{YM.X}$. In contrast, the ESM approaches almost never declared substantial mediation when it was not present at small sample sizes. All approaches rarely declared full mediation when $b^*_{MX} b^*_{YM.X} = 0$, but the ESM approaches had the best performance. The results for the K procedure fall in between the BK/SEM procedures and the ESM procedures, namely that the risk of falsely declaring there is a large proportion of variability attributed to the indirect effect with small sample sizes was small, but not zero. At larger sample sizes, all of the approaches rarely found full or substantial mediation when it did not exist and had good power for detecting full or substantial mediation when it was present. However, it is important to note that the power for the BK and SEM approaches have a ceiling of $1 - \alpha$, because they are relying on non-rejection of the null hypothesis H_0: $b^*_{YX.M} = 0$, whereas the power for the ESM and K approaches will approach unity as sample sizes increase. The choice between the K procedure or the ESM procedures with small sample sizes comes down to whether power is more important (in which case the K procedure is preferred) or whether strict control over the probability of falsely concluding full/substantial mediation is more important (in which case the ESM procedures are preferred).

There are a couple of limitations of the study/method that are necessary to highlight. First, like with all simulation studies, the results are specific to the conditions investigated. However, we are confident that the general conclusions of the study are robust given the variety of possible conditions presented. Second, as discussed earlier, the selection of the appropriate equivalence interval is a key part of equivalence testing and its importance cannot be overlooked. It is a difficult task given that we are exploring regression parameters whose distributions are less than straightforward. Bootstrapping was used to obtain a standard error for the difference between the coefficients, however what constitutes the smallest meaningful difference (δ) is still up for debate. In our simulation study, we used standardized variables in order to provide some context to the process of selecting an appropriate δ, however, as we stated earlier, an appropriate magnitude will also de-

pend on the nature of the study, etc. As equivalence testing becomes more mainstream it is hoped that much more research will be conducted into appropriate bounds for these types of problems (see Beribisky et al., 2019).

It cannot be overstated that the BK and SEM approaches to testing for full mediation violate the underlying principles of null hypothesis testing because the research hypothesis is aligned with the null hypothesis (H_0: $b^*_{YX.M}$ = 0), rather than with the alternate hypothesis. The ESM procedures, on the other hand, focus on the concept of substantial mediation and the research hypothesis is correctly aligned with the alternate hypothesis. The ESM approaches evaluate whether the direct effect of Y on X, controlling for M ($b_{YX.M}$) is neglible (and require that either the standardized total effect, β_{YX}, or the standardized indirect effect, $\beta_{MX}\beta_{YM.X}$, are non-negligible). Given that the focus is on the direct effect (i.e., the equivalence of the indirect and total effects), there needs to be variability in the raw relationship to mediate.

There are a couple of other important points to make in terms of framing our study within the larger analytic framework. First, although we focus on the simple three variable mediation model, the concept of substantial mediation could also be tested with modern mediation models including longitudinal models, causal models, etc. (MacKinnon, Fairchild, & Fritz, 2007; Preacher, 2015; Selig & Preacher, 2009). On a related note, it could be possible that multiple mediators combine to account for a substantial portion of the relationship between X and Y, and this situation should be examined in future research. Second, it is now well known that null hypothesis significance testing on its own cannot provide a complete picture of the relationship among variables and thus it is imperative for researchers to provide effect sizes (and confidence intervals for the effect sizes) to accompany their significance tests. The methods presented in this paper are based on null hypothesis testing, and accordingly should always be accompanied by an associated effect size measure. To date, there has not been consensus regarding the best way to quantify effect sizes for mediation (Preacher & Kelley, 2011; Wen & Fan, 2015). However, in the illustrative example, in addition to the proportion of the total effect mediated, we chose to use Lachowicz et al.'s (2018) recent recommendation of the squared completely standardized indirect effect (also interpreted as the variance in the outcome explained jointly by the predictor and mediator as a viable effect size measure). Providing multiple measures of effect size is usually very informative, and this is recommended given the current state of effect size reporting for mediation.

To summarize, researchers frequently test whether a particular mediator explains a substantial portion of the variability in a relationship and thus it is important that an appropriate test of this phenomenon be available. Traditional approaches for assessing full mediation, such as those using the Baron and Kenny (1986) and James and Brett (1984) models, do not suitably address this research question. The results of this study highlight that the best way to assess the question of full/substantial mediation is to either use a method based on effect size, such as that proposed by Kenny (2020), or the equivalence testing based ESM procedures proposed in this paper.

Authors' note

References

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. doi:10.1037/0022-3514.51.6.1173

Beribisky, N., Davidson, H., & Cribbie, R. A. (2019). Exploring perceptions of meaningfulness in visual representations of bivariate relationships. *PeerJ*, *7*, e6853. doi:10.7717/peerj.6853

Cappelleri, J. C., & Bushmakin, A. G. (2014). Interpretation of patient-reported outcomes. *Statistical Methods in Medical Research*, *23*(5), 460–483. doi:10.1177/0962280213476377

Caro, D. H. (2015). Causal mediation in educational research: An illustration using international assessment data. *Journal of Research on Educational Effectiveness*, *8*(4), 577–597. doi:10.1080/19345747.2015.1086913

Caron, P.-O., & Valois, P. (2018). A computational description of simple mediation analysis. *The Quantitative Methods for Psychology*, *14*(2), 147–158. doi:10.20982/tqmp.14.2.p147

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

Cribbie, R. A., Gruman, J. A., & Arpin-Cribbie, C. A. (2004). Recommendations for applying tests of equivalence. *Journal of Clinical Psychology*, *60*(1), 1–10. doi:10.1002/jclp.10217

Cribbie, R. A., Ragoonanan, C., & Counsell, A. (2016). Testing for negligible interaction: A coherent and robust approach. *British Journal of Mathematical and Statistical Psychology*, *69*(2), 159–174. doi:10.1111/bmsp.12066

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, *51*(1), 115–134. doi:10.1037/0022-0167.51.1.115

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

Goertzen, J., & Cribbie, R. A. (2010). Detecting a lack of association: An equivalence testing approach. *British Journal of Mathematical and Statistical Psychology, 63*(3), 527–537. doi:10.1348/000711009X475853

Hadlandsmyth, K., & Vowles, K. E. (2009). Does depression mediate the relation between fatigue severity and disability in chronic fatigue syndrome sufferers? *Journal of Psychosomatic Research*, *66*(1), 31–35. doi:10.1016/j.jpsychores.2008.08.002

Hayes, A. F. (2009). Beyond baron and kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420. doi:10.1080/03637750903310360

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: Process versus structural equation modeling. *Australasian Marketing Journal*, *25*(1), 76–81. doi:10.1016/j.ausmj.2017.02.001

Holmbeck, G. N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology*, *27*(1), 87–96. doi:10.1093/jpepsy/27.1.87

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation. *Journal of Consumer Psychology, 17*(2), 140–154. doi:10.1016/S1057-7408(07)70020-7

James, L. R., & Brett, J. M. (1984). Mediators, moderators and tests for mediation. *Journal of Applied Psychology*, *69*(2), 307–321. doi:10.1037/0021-9010.69.2.307

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods*, *9*(2), 233–244. doi:10.1177/1094428105285144

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*(5), 602–619. doi:10.1177/0193841X8100500502

Keele, L., Tingley, D., & Yamamoto, T. (2015). Identifying mechanisms behind policy interventions via causal mediation analysis. *Journal of Policy Analysis and Management*, *34*(4), 937–963. doi:10.1002/pam.21853

Kelley, K. (2007). Methods for the behavioral, educational, and social sciences: An r package. *Behavior Research Methods*, *39*(4), 979–984. doi:10.3758/BF03192993

Kenny, D. A. (2020). Mediation. Retrieved January 1, 2020, from http://davidakenny.net/cm/mediate.htm

Kenny, D. A., & Judd, C. M. (2013). Power anomalies in testing mediation. *Psychological Science*, *25*(2), 334–339. doi:10.1177/0956797613502676

Kukihara, H., Yamawaki, N., Ando, M., Nishio, M., Koga, K., Kimura, H., & Matsuda, T. (2020). Effects of exercise and mindfulness-based yoga programs on promotion of resilience and mental health of older adults in japan: A randomized controlled trial. *Psychology*, *11*(2), 285–298. doi:10.4236/psych.2020.112018

Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, *23*(2), 244–261. doi:10.1037/met0000165

Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In P. F. Lazardsfeld & M. Rosenberg (Eds.), *The language of social research: A reader in the methodology of social research* (pp. 115–125). Glencoe, Il.: Free Press.

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., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, *1*(4), 173–186. doi: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 the significance of the mediated effect. *Psychological Methods*, *7*(1), 83–104. doi:10.1037/1082-989X.7.1.83

Namazi, M., & Namazi, N. R. (2016). Conceptual analysis of moderator and mediator variables in business research. *Procedia Economics and Finance*, *36*, 540–554. doi:10.1016/S2212-5671(16)30064-8

Pek, J., & Hoyle, R. H. (2016). On the (in) validity of tests of simple mediation: Threats and solutions. *Social and Personality Psychology Compass*, *10*(3), 150–163. doi:10.1111/spc3.12237

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*, 825–852. doi:10.1146/annurev-psych-010814-015258

Preacher, K. J., & Hayes, A. F. (2004). Spss and sas procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731. doi:10.3758/BF03206553

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

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for com-

municating indirect effects. *Psychological Methods*, *16*(2), 93–115. doi:10.1037/a0022658

R Development Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ritt-Olson, A., Unger, J., Valente, T., Nezami, E., Chou, C.-P., Trinidad, D., . . . Anderson-Johnson, C. (2005). Exploring peers as a mediator of the association between depression and smoking in young adolescents. *Substance Use and Misuse*, *40*(1), 77–98. doi:10.1081/JA-200030505

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin*, *113*(3), 553–565. doi:10.1037/0033-2909.113.3.553

Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling and more. *Journal of Statistical Software*, *48*, 1–36. doi: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*, 359–371. doi:10.1111/j.1751-9004.2011.00355.x

Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, *15*(6), 657–680. doi:10.1007/BF01068419

Seaman, M. A., & Serlin, R. C. (1998). Equivalence confidence intervals for two-group comparisons of means. *Psychological Methods*, *3*(4), 403–411. doi:10.1037/1082-989X.3.4.403

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2-3), 144–164. doi:10.1080/15427600902911247

Sella, F., Sader, E., Lolliot, S., & Cohen Kadosh, R. (2016). Basic and advanced numerical performances relate to mathematical expertise but are fully mediated by visuospatial skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(9), 1458–1472. doi:10.1037/xlm0000249

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

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290–312). San Francisco: Jossey-Bass.

Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological Methodology*, *16*, 159–186. doi:10.2307/270922

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251. doi:10.1037/0033-2909.87.2.245

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods*, *6*(4), 371–386. doi:10.1037/1082-989X.6.4.371

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority (2nd ed.)* Boca Raton: CRC Press.

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

Citation

Tables 1 to 5 follow.

Table 1 ■ Proportion of indications of full mediation for the BK, SEM, and K procedures and substantial mediation for the ESM procedures, $b^*_{M.X} b^*_{Y.M.X} = ab = 0$. For $a > b$: $a = .50$, $b = 0$; $a = b = 0$; $a = b$: $a = b = 0$; $a < b$: $a = 0$, $b = .50$.

N	$b^*_{Y.X.M}$	BK			SEM			K			ESM_{ind}			ESM_{tot}		
		$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$
50	0	.003	.000	.014	.031	.001	.024	.000	.000	.003	.000	.000	.000	.000	.000	.000
	.05	.001	.000	.010	.028	.002	.027	.000	.000	.007	.000	.000	.000	.000	.000	.000
	.25	.004	.001	.010	.095	.002	.012	.001	.000	.000	.000	.000	.000	.000	.000	.000
100	0	.002	.000	.008	.031	.001	.024	.000	.000	.001	.000	.000	.000	.000	.000	.000
	.05	.004	.000	.011	.034	.000	.031	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.25	.014	.000	.003	.001	.000	.004	.000	.000	.000	.000	.000	.000	.001	.000	.000
200	0	.002	.000	.010	.022	.001	.023	.000	.000	.000	.000	.000	.000	.001	.000	.001
	.05	.003	.001	.012	.058	.003	.026	.000	.000	.000	.000	.000	.000	.001	.000	.001
	.25	.010	.000	.001	.060	.001	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
350	0	.000	.000	.009	.023	.000	.027	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.05	.001	.000	.007	.045	.000	.020	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.25	.004	.000	.000	.008	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
500	0	.000	.000	.013	.020	.000	.020	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.05	.004	.000	.010	.056	.002	.018	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.25	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
1500	0	.001	.000	.010	.030	.002	.029	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.05	.013	.000	.005	.017	.000	.007	.000	.000	.000	.000	.000	.000	.000	.000	.000
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note. N = sample size; $b_{Y.X.M}$ = coefficient for the direct effect after controlling for the mediator; BK= Baron and Kenny (1986) approach to assessing full mediation; SEM = James and Brett (1984) structural equation modeling based approach to assessing full mediation; K = Kenny full mediation approach using the proportion mediated; ESM_{ind} = Equivalence test of substantial mediation, conditional upon minimum $a \times b$ coefficient; ESM_{tot} = Equivalence test of substantial mediation conditional upon minimum c coefficient.

Table 2 ■ Proportion of indications of full mediation for the BK, SEM, and K procedures and substantial mediation for the ESM procedures, $b^*_{M.X} b^*_{Y.M.X}$ = ab = .09. For a > b: a = .519, b = .173; a = b: a = b = .3; a < b: a = .173, b = .519.

N	$b^*_{Y.X.M}$	BK			SEM			K			ESM_{ind}			ESM_{tot}		
		a > b	a = b	a < b	a > b	a = b	a < b	a > b	a = b	a < b	a > b	a = b	a < b	a > b	a = b	a < b
50	0	.014	.035	.046	.240	.364	.226	.029	.015	.035	.000	.001	.000	.000	.001	.001
	.05	.024	.063	.069	.293	.375	.236	.025	.022	.026	.000	.000	.000	.000	.001	.001
	.25	.105	.123	.079	.404	.275	.105	.037	.008	.012	.000	.000	.001	.000	.000	.001
100	0	.052	.047	.088	.404	.737	.379	.014	.005	.015	.000	.003	.008	.000	.003	.012
	.05	.080	.163	.144	.516	.731	.374	.013	.007	.025	.000	.000	.011	.000	.003	.016
	.25	.179	.180	.066	.324	.213	.075	.016	.002	.002	.000	.000	.000	.000	.000	.000
200	0	.131	.237	.206	.647	.920	.659	.001	.002	.005	.009	.008	.029	.015	.018	.060
	.05	.248	.390	.339	.722	.884	.590	.006	.004	.003	.010	.008	.029	.016	.023	.064
	.25	.086	.039	.004	.094	.040	.004	.001	.000	.000	.001	.000	.000	.001	.000	.000
350	0	.319	.362	.344	.860	.953	.858	.002	.000	.000	.017	.004	.025	.037	.035	.075
	.05	.539	.608	.532	.852	.842	.723	.002	.000	.002	.003	.005	.024	.066	.084	.161
	.25	.012	.001	.000	.014	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
500	0	.459	.479	.491	.923	.951	.941	.000	.000	.000	.005	.003	.008	.039	.033	.062
	.05	.717	.671	.641	.855	.805	.738	.001	.000	.000	.017	.002	.005	.101	.149	.222
	.25	.002	.000	.000	.002	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.000
1500	0	.920	.918	.923	.962	.952	.958	.000	.000	.000	.000	.000	.000	.006	.009	.008
	.05	.622	.509	.377	.621	.509	.376	.000	.000	.000	.000	.000	.000	.279	.331	.356
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note. N = sample size; $b_{Y.X.M}$ = coefficient for the direct effect after controlling for the mediator; BK= Baron and Kenny (1986) approach to assessing full mediation; SEM = James and Brett (1984) structural equation modeling based approach to assessing full mediation; K = Kenny full mediation approach using the proportion mediated; ESM_{ind} = Equivalence test of substantial mediation, conditional upon minimum $a \times b$ coefficient; ESM_{tot} = Equivalence test of substantial mediation conditional upon minimum c coefficient.

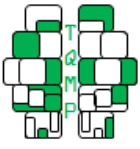**Table 3 ■** Proportion of indications of full mediation for the BK, SEM, and K procedures and substantial mediation for the ESM procedures, $b^*_{M.X} b^*_{Y.M.X} = ab = .25$. For $a > b$: $a = .709, b = .353; a = b: a = b = .5; a < b: a = .353, b = .709$.

N	$b^*_{Y.X.M}$	BK			SEM			K			ESM_{ind}			ESM_{tot}		
		$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$
50	0	.184	.381	.365	.719	.900	.677	.305	.340	.377	.000	.001	.014	.000	.001	.016
	.05	.270	.483	.514	.784	.892	.705	.347	.314	.360	.000	.001	.019	.000	.001	.022
	.25	.430	.452	.148	.634	.465	.149	.144	.079	.018	.000	.001	.001	.000	.001	.001
100	0	.514	.697	.696	.925	.941	.914	.367	.451	.498	.001	.029	.241	.001	.029	.249
	.05	.625	.795	.767	.917	.918	.852	.394	.413	.426	.000	.024	.216	.000	.024	.224
	.25	.442	.205	.021	.440	.196	.020	.087	.036	.002	.000	.004	.002	.000	.004	.002
200	0	.891	.914	.921	.957	.944	.950	.494	.562	.664	.048	.341	.749	.048	.341	.731
	.05	.893	.903	.826	.919	.902	.831	.477	.482	.530	.045	.294	.576	.045	.294	.580
	.25	.166	.025	.001	.162	.025	.001	.020	.003	.001	.001	.000	.001	.001	.000	.001
350	0	.922	.921	.918	.923	.921	.916	.593	.686	.752	.331	.740	.945	.331	.736	.925
	.05	.873	.851	.753	.872	.848	.751	.550	.580	.606	.303	.607	.809	.303	.607	.808
	.25	.028	.001	.000	.026	.001	.000	.006	.000	.000	.001	.000	.000	.001	.000	.000
500	0	.931	.938	.936	.929	.937	.935	.668	.773	.820	.583	.900	.992	.584	.898	.988
	.05	.875	.789	.697	.874	.786	.692	.570	.580	.663	.504	.718	.935	.504	.717	.934
	.25	.008	.000	.000	.008	.000	.000	.006	.000	.000	.001	.000	.000	.001	.000	.000
1500	0	.950	.947	.944	.949	.946	.944	.931	.965	.979	.994	1.00	1.00	.994	1.00	1.00
	.05	.684	.520	.227	.684	.520	.227	.631	.695	.700	.901	1.00	1.00	.901	1.00	1.00
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note. N = sample size; $b_{Y.X.M}$ = coefficient for the direct effect after controlling for the mediator; BK= Baron and Kenny (1986) approach to assessing full mediation; SEM = James and Brett (1984) structural equation modeling based approach to assessing full mediation; K = Kenny full mediation approach using the proportion mediated; ESM_{ind} = Equivalence test of substantial mediation, conditional upon minimum $a \times b$ coefficient; ESM_{tot} = Equivalence test of substantial mediation conditional upon minimum c coefficient.
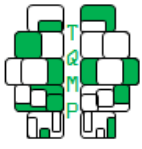
Table 4 ■ Proportion of indications of full mediation for the BK, SEM, and K procedures and substantial mediation for the ESM procedures, $b^*_{M.X} b^*_{Y.M.X} = ab = .36$. For $a > b$: $a = .735, b = .490$; $a = b$: $a = b = .6$; $a < b$: $a = .490, b = .735$.

N	$b^*_{Y.X.M}$	BK			SEM			K			ESM_{ind}			ESM_{tot}		
		$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$
50	0	.552	.699	.691	.919	.932	.906	.564	.576	.646	.000	.000	.019	.000	.000	.019
	.05	.627	.807	.776	.921	.920	.889	.530	.549	.579	.000	.001	.028	.000	.001	.028
	.25	.641	.421	.105	.634	.401	.099	.202	.098	.041	.000	.002	.001	.000	.002	.001
100	0	.904	.924	.930	.954	.951	.958	.702	.746	.817	.000	.026	.219	.000	.026	.219
	.05	.903	.909	.884	.926	.916	.887	.591	.612	.666	.001	.015	.229	.001	.015	.229
	.25	.353	.153	.006	.343	.143	.005	.102	.038	.002	.000	.001	.001	.000	.001	.001
200	0	.944	.959	.950	.944	.958	.950	.781	.861	.921	.048	.354	.694	.048	.354	.694
	.05	.910	.878	.835	.908	.877	.833	.649	.702	.759	.066	.313	.574	.066	.313	.574
	.25	.091	.006	.000	.088	.005	.000	.035	.004	.001	.001	.000	.000	.001	.000	.000
350	0	.949	.959	.946	.946	.959	.944	.886	.930	.972	.408	.756	.940	.408	.756	.940
	.05	.888	.850	.744	.888	.848	.740	.700	.756	.810	.365	.606	.779	.365	.606	.779
	.25	.007	.000	.000	.006	.000	.000	.005	.000	.000	.000	.000	.000	.000	.000	.000
500	0	.935	.961	.949	.934	.959	.948	.912	.976	.992	.656	.905	.992	.656	.905	.992
	.05	.856	.801	.669	.856	.800	.666	.729	.802	.857	.534	.735	.897	.534	.735	.897
	.25	.001	.000	.000	.001	.000	.000	.002	.000	.000	.000	.000	.000	.000	.000	.000
1500	0	.950	.957	.946	.950	.957	.946	.990	.999	1.00	.996	1.00	1.00	.996	1.00	1.00
	.05	.668	.489	.260	.668	.489	.258	.871	.924	.980	.921	.994	.999	.921	.994	.999
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note. N = sample size; $b_{Y.X.M}$ = coefficient for the direct effect after controlling for the mediator; BK = Baron and Kenny (1986) approach to assessing full mediation; SEM = James and Brett (1984) structural equation modeling based approach to assessing full mediation; K = Kenny full mediation approach using the proportion mediated; ESM_{ind} = Equivalence test of substantial mediation, conditional upon minimum $a \times b$ coefficient; ESM_{tot} = Equivalence test of substantial mediation conditional upon minimum c coefficient.
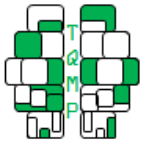
Table 5 ■ Proportion of indications of full mediation for the BK, SEM, and K procedures and substantial mediation for the ESM procedures, $b^*_{M.X} b^*_{Y.M.X} = ab = .49$. For $a > b$: $a = .808$, $b = .606$; $a = b$: $a = b = .7$; $a < b$: $a = .606$, $b = .808$.

N	$b^*_{Y.X.M}$	BK			SEM			K			ESM_{ind}			ESM_{tot}		
		$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$	$a > b$	$a = b$	$a < b$
50	0	.787	.924	.920	.937	.940	.950	.695	.722	.833	.000	.000	.002	.000	.000	.002
	.05	.837	.937	.906	.926	.933	.914	.626	.682	.752	.000	.003	.038	.000	.003	.038
	.25	.577	.222	.000	.546	.203	.000	.207	.108	.007	.000	.002	.000	.000	.002	.000
100	0	.941	.961	.945	.944	.958	.941	.766	.857	.931	.000	.016	.282	.000	.016	.282
	.05	.934	.919	.908	.930	.910	.904	.687	.750	.832	.001	.036	.321	.001	.036	.321
	.25	.295	.038	.000	.288	.032	.000	.145	.047	.000	.000	.001	.000	.000	.001	.000
200	0	.948	.946	.956	.945	.943	.953	.844	.929	.985	.021	.365	.794	.021	.365	.794
	.05	.919	.892	.824	.916	.888	.817	.759	.827	.903	.043	.323	.649	.043	.323	.649
	.25	.043	.001	.000	.040	.001	.000	.051	.005	.000	.000	.000	.000	.000	.000	.000
350	0	.956	.956	.950	.955	.956	.949	.938	.977	.995	.324	.754	.959	.324	.754	.959
	.05	.894	.827	.732	.890	.825	.728	.807	.893	.966	.312	.601	.854	.312	.601	.854
	.25	.002	.000	.000	.002	.000	.000	.015	.001	.000	.000	.000	.000	.000	.000	.000
500	0	.940	.946	.952	.939	.945	.952	.958	.991	1.00	.590	.910	.997	.590	.910	.997
	.05	.862	.775	.609	.860	.773	.608	.859	.933	.979	.490	.752	.936	.490	.752	.936
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
1500	0	.958	.954	.945	.958	.953	.945	1.00	1.00	1.00	.994	1.00	1.00	.994	1.00	1.00
	.05	.684	.494	.205	.683	.492	.204	.975	.997	1.00	.905	.993	1.00	.905	.993	1.00
	.25	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note. N = sample size; $b_{Y.X.M}$ = coefficient for the direct effect after controlling for the mediator; BK= Baron and Kenny (1986) approach to assessing full mediation; SEM = James and Brett (1984) structural equation modeling based approach to assessing full mediation; K = Kenny full mediation approach using the proportion mediated; ESM_{ind} = Equivalence test of substantial mediation, conditional upon minimum $a \times b$ coefficient; ESM_{tot} = Equivalence test of substantial mediation conditional upon minimum c coefficient.