

Comparison of methods for constructing confidence intervals of standardized indirect effects

MIKE W.-L. CHEUNG

National University of Singapore, Kent Ridge, Singapore

Mediation models are often used as a means to explain the psychological mechanisms between an independent and a dependent variable in the behavioral and social sciences. A major limitation of the unstandardized indirect effect calculated from raw scores is that it cannot be interpreted as an effect-size measure. In contrast, the standardized indirect effect calculated from standardized scores can be a good candidate as a measure of effect size because it is scale invariant. In the present article, 11 methods for constructing the confidence intervals (CIs) of the standardized indirect effects were evaluated via a computer simulation. These included six Wald CIs, three bootstrap CIs, one likelihood-based CI, and the PRODCLIN CI. The results consistently showed that the percentile bootstrap, the bias-corrected bootstrap, and the likelihood-based approaches had the best coverage probability. Mplus, LISREL, and Mx syntax were included to facilitate the use of these preferred methods in applied settings. Future issues on the use of the standardized indirect effects are discussed.

A mediator is a variable that explains the psychological mechanism between an independent variable and a dependent variable (see, e.g., Baron & Kenny, 1986; James & Brett, 1984; Judd & Kenny, 1981; MacKinnon, 2008; MacKinnon, Fairchild, & Fritz, 2007).¹ A simple literature review shows that Baron and Kenny's seminal article has been cited more than ten thousand times. Their article is also the most cited in the history of the *Journal of Personality & Social Psychology* (Quinones-Vidal, Lopez-Garcia, Penaranda-Ortega, & Tortosa-Gil, 2004). Although some of the citations may be related to moderators rather than to mediators, this large number still clearly indicates the importance of using mediation models in the social and behavioral sciences (see MacKinnon, 2008, for the applications of mediation models in many disciplines).

Figure 1 shows a model with one single mediator. X , M , and Y are the independent variable, the mediator, and the dependent variable, respectively. The basic idea of a mediation analysis is to decompose the total effect into two parts: the direct and the indirect effects. The direct and the indirect effects can be estimated by the product terms c' and ab , respectively, whereas the sum of them ($ab + c'$) is known as the total effect. In order to conduct a mediation analysis, both regression analysis and structural equation modeling (SEM) may be used (see, e.g., M. W.-L. Cheung, 2007; James, Mulaik, & Brett, 2006; MacKinnon, 2008; Preacher & Hayes, 2004, 2008a).

After obtaining an estimate on the indirect effect, it is also important for one to report its precision. The precision of an estimated indirect effect can be in the form of either a standard error (SE ; Aroian, 1944; Goodman, 1960; Sobel,

1982, 1986) or a confidence interval (CI; Bollen & Stine, 1990; M. W.-L. Cheung, 2007, 2009; MacKinnon, Lockwood, & Williams, 2004; Shrout & Bolger, 2002; Williams & MacKinnon, 2008). The present study aimed to focus mainly on the CI, even though many of the techniques discussed can be directly applied to SE .

There were three main objectives in the present study. First, it emphasized the importance of using effect size and its CI in a mediation analysis. Such a practice is consistent with the current statistical reform in psychology and in other disciplines in the social sciences (see, e.g., Cumming et al., 2007; Wilkinson et al., 1999). Effect size is also required in power analysis, along with sample-size estimation and defining conditions for computer simulations (see, e.g., the studies conducted by MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002).

On the basis of a simulation study, Fritz and MacKinnon (2007) presented sample sizes required to test small, medium, and large effect sizes in a mediation analysis. In their study, and in all other simulation studies conducted by MacKinnon and his colleagues (e.g., MacKinnon et al., 2002), the model specification was set in such a way that the error variances of the mediator and the dependent variable were both 1.0.² In other words, the regression coefficients for small (.14), medium (.39), and large (.59) effects are interpreted under the assumption that the error variances of the mediator and the dependent variable are both 1.0. This makes it difficult to interpret the strength of the indirect effect, since the error variances are seldom equal to 1.0 in applied settings. The present study suggested using a standardized indirect effect in a mediation analysis. It will

M. W.-L. Cheung, mikewlcheung@nus.edu.sg

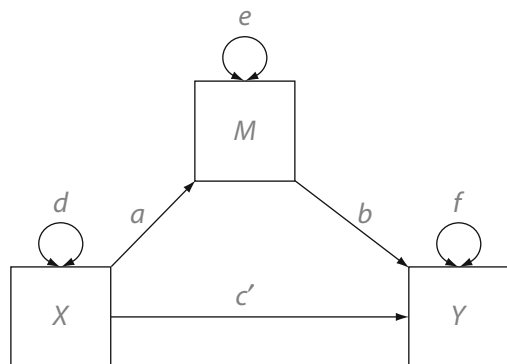


Figure 1. A mediating model with unstandardized variables.

be shown that a standardized indirect effect is much more interpretable.

The second objective was to review the correct procedures for constructing CIs of the standardized indirect effect. It is well known that statistical tests based on correlation and covariance matrices can be different (see, e.g., Bollen, 1989; Cudeck, 1989; Wansbeek & Meijer, 2000). MacKinnon (2008) also explicitly stated that “[u]sing the [formulas derived for unstandardized mediation analysis] to test the significance of the product of standardized \hat{a} and \hat{b} regression coefficients (based on unstandardized standard errors) can lead to misleading results” (p. 74). Even worse, some researchers may wrongly apply bootstrap procedures to the standardized data. This approach is labeled as *naive bootstrap CI* in the present article. The present study showed the problems with this naive bootstrap method and the correct way to obtain bootstrap CIs of the standardized indirect effect.

The final objective was to empirically evaluate 11 methods for constructing CIs of the standardized indirect effect. These methods included 6 Wald CIs based on *SEs*, 3 bootstrap CIs, 1 likelihood-based CI, and the PRODCLIN (distribution of the product confidence limits for indirect effects; MacKinnon, Fritz, Williams, & Lockwood, 2007) CI. Although there are several simulation studies on the performance of these methods in testing the unstandardized indirect effect, it is not clear how good these methods are in testing the standardized indirect effect. The present study filled this research gap by providing some empirical findings via a computer simulation.

The present article is organized as follows. In the next section, I will briefly discuss the importance of using effect size in a mediation analysis. In the subsequent section, methods for constructing CIs of the unstandardized indirect effect and the ways in which they can be extended to the standardized indirect effect will be reviewed. A computer simulation will then be presented to evaluate the empirical coverage probabilities of these methods. Finally, future directions for and issues related to the standardized indirect effect will be discussed.

Importance of Reporting Effect Size in Mediation Analysis

Conventionally, researchers focus on the significance of the indirect effect. Since the magnitude of the indirect effect depends on the scales of both the independent and

the dependent variables (see Equation A1 in Appendix A), the unstandardized indirect effect calculated from raw scores cannot be interpreted as a measure of the effect size. MacKinnon (2008) commented that “[t]he value of the [unstandardized] mediated effect is more interpretable if the unit of measurement of the dependent variable involved is clear” (p. 79).

However, the scaling of many variables in the social and behavioral sciences is arbitrary. This creates serious problems with interpreting the magnitude of the unstandardized indirect effect. Besides the significance of the indirect effect, researchers cannot say anything about the strength of the indirect effect. Moreover, researchers cannot compare the indirect effect found in one study with those found in other studies. Researchers may use complete or partial mediation to qualify the type of mediation involved (Baron & Kenny, 1986; James et al., 2006). However, doing this still will not lead to the quantification of the indirect effect.

In the social and behavioral sciences, the calculation of effect sizes and their CIs is usually preferred over null hypothesis significance testing (see, e.g., American Psychological Association, 2001; Cohen, 1994; Harlow, Mulaik, & Steiger, 1997; Wilkinson et al., 1999). The *Publication Manual of the American Psychological Association* (2001) explicitly states that “it is almost always necessary to include some index of effect size or strength of relationship in your Results section” (p. 25). An effect size provides a scale-free measure of the strength of association that can be compared and synthesized across studies.

MacKinnon, Warsi, and Dwyer (1995; see also MacKinnon, 2008b) compared several potential measures of effect size in the context of mediation analysis. Two of them were the ratio mediated (ab/c' ; Sobel, 1982) and the proportion mediated [$ab/(ab + c')$; Alwin & Hauser, 1975]. The ratio mediated is interpreted as the ratio of the indirect effect to the direct effect, whereas the proportion mediated is interpreted as the proportion of the effect from X to Y that can be accounted for by the mediator, M . However, both interpretations become problematic when there is an inconsistent mediation (MacKinnon, Krull, & Lockwood, 2000). When inconsistent mediation occurs, the signs between ab and c' are different, and the terms ab/c' and $ab/(ab + c')$ can be negative. Moreover, MacKinnon et al. (1995) found that these measures are not very stable unless the sample size is at least 500. Another limitation is that these two measures do not indicate the effect of the mediation; they merely show the ratio (or percentage) of the indirect effect to the direct effect (or the total effect).

A more direct measure of the effect size is the standardized indirect effect calculated from the standardized scores (see, e.g., Bobko & Rieck, 1980; Preacher & Hayes, 2008b; Raykov, Brennan, Reinhardt, & Horowitz, 2008). There are several advantages to using a standardized indirect effect. First, it is interpreted in a manner similar to that for the unstandardized indirect effect; that is, it is interpreted as the effect of an independent variable on a dependent variable via a mediator after controlling for the direct effect when all variables are standardized. Second, the standardized total effect always equals the sum of the standardized indirect effect and the standardized direct ef-

fect. Researchers do not need to learn new concepts before interpreting a standardized indirect effect.

Third, the standardized indirect effect is usually within the range of -1 to $+1$, given that there is no inconsistent mediation. Statistically speaking, the standardized indirect effect can be larger (or smaller) than 1 (or -1) (Jöreskog, 1999), although this is not likely. This does not affect the use of the standardized indirect effect as a measure of effect size, however. Being bounded by a certain value is not a requirement for an effect-size measure. For example, there is no boundary on the odds ratio and on the standardized mean difference. In fact, both of them are important measures of effect size in the behavioral sciences. Since the standardized indirect effect indicates both the direction of an influence and the strength of an association, it can serve as a measure of effect size in a mediation analysis.

Similar concepts of the standardized indirect effect have been widely applied in simulation studies on unstandardized indirect effects in order to define the population of effect sizes (see, e.g., G. W. Cheung & Lau, 2008; M. W.-L. Cheung, 2007; Fritz & MacKinnon, 2007; MacKinnon, Fairchild, & Fritz, 2007; MacKinnon et al., 2002; MacKinnon et al., 2004; Taylor, MacKinnon, & Tein, 2008) (see note 2). The use of the standardized indirect effect also facilitates the comparison between indirect effects across studies and a meta-analysis (Hedges & Olkin, 1985; Hunter & Schmidt, 2004). Since studies are likely different in terms of measures, scales, and samples, standardized scores are preferred in a meta-analysis (see Hunter & Hamilton, 2002, for the arguments on this matter). Therefore, the standardized indirect effect is not only useful as an effect-size measure in primary studies; it is also useful for future meta-analyses that cumulate scientific knowledge regarding the mediating effect on a particular topic.

Indeed, the task of estimating the standardized indirect effect is not a new one to many SEM users. For example, popular SEM packages such as LISREL (Jöreskog & Sörbom, 1996) and Mplus (Muthén & Muthén, 2007) provide commands to obtain standardized indirect effects. However, estimating the CIs of the standardized indirect effect has been investigated less in the literature.³

Constructing Confidence Intervals of Unstandardized and Standardized Indirect Effects

Standardizing Raw Scores Before a Mediation Analysis

Before discussing methods of constructing CIs of the standardized indirect effect, I will first address issues related to the naive bootstrap CI. Intuitively, one may simply want to standardize the data and then conduct a mediation analysis on the standardized data. The CIs obtained for the standardized data may then be used as the CIs for the standardized indirect effect.

Although the present approach looks appealing because the existing methods for unstandardized indirect effects can be applied directly to standardized indirect effects without any modification, several issues should be noted. First, in standardizing the data prior to conducting a mediation analysis, the assumption is that the estimates of both V_X and V_Y

(variances of X and Y) are fixed. That is, variability in estimating V_X and V_Y is not incorporated into the calculations of the CI of the standardized indirect effect. The Wald CIs obtained by this approach are the same as the one obtained by using Equation 7, which will be discussed later.

Second—and most important—bootstrap procedures should not be applied directly to the standardized data to obtain the bootstrap CIs. The reason is that the variances of the bootstrapped samples may be different from 1.0 , even when the data have been standardized prior to the analysis. Therefore, the CIs constructed from the bootstrapped samples cannot be qualified as the CIs of the standardized indirect effect. The correct procedure is to standardize each bootstrapped sample before calculating the standardized indirect effect. Doing so ensures that the variances of the bootstrapped samples are exactly 1.0 .

Third, standardizing the data prior to the analysis may not be appropriate for more complicated mediational models—for example, those with latent variables and with specific or intermediate mediators. This is similar to the analysis of correlation matrices by using covariance structure models in the context of SEM (Cudeck, 1989). The SE s and CIs of the parameter estimates may be incorrect. A model-based standardization (see, e.g., Jöreskog, Sörbom, Du Toit, & Du Toit, 1999), which will be introduced later, is recommended to obtain the standardized coefficients.

Methods of Constructing Confidence Intervals for the Unstandardized Indirect Effect

In this section, I will summarize several methods of obtaining CIs for the unstandardized indirect effect. Then I will extend these methods to the case of the standardized indirect effect. In Figure 1, the unstandardized indirect effect is estimated by

$$I_{\text{unstand.}} = ab. \quad (1)$$

On the basis of the multivariate delta method with a first-order Taylor series approximation, Sobel (1982) derived the SE of the unstandardized indirect effect as follows:

$$SE_{\text{Sobel}} = \sqrt{a^2 SE_b^2 + b^2 SE_a^2}, \quad (2)$$

where SE_a and SE_b are the estimated asymptotic standard errors of a and b , respectively (see also MacKinnon, 2008, for the derivations).⁴

There are other, alternative estimators of the SE of the unstandardized indirect effect. Aroian (1944) and Goodman (1960) estimated the SE of the unstandardized indirect effect by using the multivariate delta method with a first- and second-order Taylor series approximation and with the distribution of the product term of two random variables. Their proposed estimators were

$$SE_{\text{Aroian}} = \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + SE_a^2 SE_b^2}, \quad (3)$$

and

$$SE_{\text{Goodman}} = \sqrt{a^2 SE_b^2 + b^2 SE_a^2 - SE_a^2 SE_b^2}. \quad (4)$$

The indirect effect divided by its estimated SE —that is, $z = ab/SE$ —approximately follows a standard normal distribu-

tion in large samples. If the absolute value of the z score is larger than 1.96, it is statistically significant at .05.

By assuming that the unstandardized indirect effect is approximately normally distributed, a 95% Wald CI can be constructed by

$$ab \pm 1.96SE, \quad (5)$$

where SE is the estimated SE based on Equations 2–4. MacKinnon et al. (2002) compared the empirical performance of these estimators via a computer simulation. They found that these tests had low Type I error rates and very low statistical power. Moreover, MacKinnon et al. (2002) and MacKinnon et al. (2004) found that SE_{Goodman} was often undefined. Therefore, SE_{Goodman} will not be further considered in the present article.

Since the distribution of the indirect effect ab is usually nonnormal except in very large sample sizes, several approaches have been suggested to test the unstandardized indirect effect. One recommended method is to use bootstrap CIs. Bootstrap methods use the empirical distribution of the statistics to approximate the theoretical distribution of the statistics (see, e.g., Bollen & Stine, 1990; MacKinnon et al., 2004; Shrout & Bolger, 2002).

Suppose that we have N data; we may take N independent draws with replacement from the original samples. The bootstrap sample is then used to construct an estimate, ab^* , of the unstandardized indirect effect. This procedure is repeated B times. On the basis of ab^* , there are several methods by which to construct bootstrap CIs. One of them is the percentile bootstrap approach. To construct a 95% percentile bootstrap CI for the unstandardized indirect effect, one may obtain the 2.5th and 97.5th percentiles on the B bootstrap replications ab^* . Several modifications may be used to improve the performance. These include the bias-corrected (BC) bootstrap CI and the bias-corrected and acceleration (BCa) bootstrap CI (see Davison & Hinkley, 1997).

Another approach for testing the unstandardized indirect effect is based on the distribution of a product. MacKinnon, Fritz, et al. (2007) presented a program called PRODCLIN that provides CI on the indirect effect. Since the PRODCLIN CI captures the non-normality of the distribution of the product between two regression coefficients, their approach performs much better than the Wald CIs in testing the indirect effect. The PRODCLIN approach can be extended to the standardized indirect effect by assuming that the variances of the variables are fixed.

An alternative approach is the use of likelihood-based CIs (M. W.-L. Cheung, 2007, 2009; Neale & Miller, 1997). To construct a $100(1 - \alpha)\%$ likelihood-based CI on a parameter—say, an unstandardized indirect effect—we move the parameter estimate (treated as varied) as far away as possible from its maximum likelihood estimate so that it is just statistically significant at the desired α level. Like the PRODCLIN approach, the likelihood-based CI captures the asymmetry on the distribution of the indirect effect. It should be noted that both PRODCLIN and likelihood-based CIs still assume that the raw data are normally distributed, whereas the bootstrap CIs do not make this assumption.

Methods for Constructing Confidence Intervals for the Standardized Indirect Effect

An unstandardized indirect effect approach. The first method for obtaining the standardized indirect effect is by standardizing the unstandardized indirect effect, using

$$I_{\text{stand.}} = \frac{V_X^{0.5}}{V_Y^{0.5}} I_{\text{unstand.}}, \quad (6)$$

where V_X and V_Y are the variances of X and Y , respectively. It should be noted that the variance of M (the mediator) is not involved in the calculations (see Appendix A).⁵

Several methods may be applied to derive the SE of $I_{\text{stand.}}$ depending on whether the estimates of V_X and V_Y are treated as fixed or random. If the sample sizes are sufficiently large, we may treat both V_X and V_Y as fixed and known (e.g., Bollen & Stine, 1990; Cohen, Cohen, West, & Aiken, 2003). Doing this is equivalent to conducting a mediation analysis on the standardized data that were mentioned before. Since both V_X and V_Y are treated as constants, the SE of the standardized indirect effect can be easily obtained by using

$$SE_{\text{stand.-fixed}} = \frac{V_X^{0.5}}{V_Y^{0.5}} SE_{\text{unstand.}}. \quad (7)$$

When the sample sizes are small, however, the estimates of V_X and V_Y may vary across samples. We have to take the variability of V_X and V_Y into account in deriving the SE of the standardized indirect effect (see, e.g., Bentler, 2007; Bentler & Lee, 1983; Bollen, 1989; Wansbeek & Meijer, 2000). By treating the estimates of ab , V_X , and V_Y as uncorrelated, the SE of the standardized indirect effect can be obtained by the multivariate delta method (see Appendix A).

$$SE_{\text{stand.-random}} = \sqrt{SE_{\text{stand.-fixed}}^2 + \left(\frac{a^2 b^2}{n-1} \right) \frac{V_X}{V_Y}}, \quad (8)$$

where n is the sample size. Since the term

$$\left(\frac{a^2 b^2}{n-1} \right) \frac{V_X}{V_Y}$$

is nonnegative, $SE_{\text{stand.-random}}$ is always larger than $SE_{\text{stand.-fixed}}$. When n gets larger and larger, $SE_{\text{stand.-random}}$ approaches $SE_{\text{stand.-fixed}}$.

Since there are two formulas (SE_{Sobel} and SE_{Aroian}) for estimating the SE of the unstandardized indirect effect, there are four possible formulas for estimating the SE of the standardized indirect effect, depending on whether we treat V_X and V_Y as fixed or random. These formulas are

$$SE_{\text{Sobel-fixed}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2}, \quad (9)$$

$$SE_{\text{Aroian-fixed}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + SE_a^2 SE_b^2}, \quad (10)$$

$$SE_{\text{Sobel-random}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + \left(\frac{a^2 b^2}{n-1} \right)}, \quad (11)$$

and

$$SE_{\text{Aroian-random}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + SE_a^2 SE_b^2 + \left(\frac{a^2 b^2}{n-1} \right)}. \quad (12)$$

A correlational approach. Another approach to obtaining the standardized indirect effect is to derive it from the correlation coefficients among X , M , and Y (Bobko & Rieck, 1980). Bobko and Rieck further proposed a formula for estimating the SE of the standardized indirect effect via the multivariate delta method (Equation 13, below), where $s = (r_{XM}^2 r_{MY} + r_{MY} - 2r_{XM} r_{XY})(1 - r_{XM}^2)^{-2}$, $t = -r_{XM}^2(1 - r_{XM}^2)^{-1}$, $u = r_{XM}(1 - r_{XM}^2)^{-1}$, $\text{var}(r_{ij})$ is the asymptotic variance of r_{ij} , and $\text{cov}(r_{ij}, r_{jk})$ is the asymptotic covariance between r_{ij} and r_{jk} .

To apply Equation 13, we have to estimate the asymptotic variance of a correlation coefficient and the asymptotic covariance between two correlation coefficients (see, e.g., Olkin & Finn, 1995; Olkin & Siotani, 1976). The asymptotic variance of a correlation coefficient r_{ij} and the asymptotic covariance between r_{ij} and r_{jk} can be approximated by

$$\text{var}(r_{ij}) = \frac{(1 - r_{ij}^2)^2}{n}$$

and

$$\text{cov}(r_{ij}, r_{jk}) = \frac{r_{ij}(1 - r_{ik}^2 - r_{jk}^2) - 0.5r_{ik}r_{jk}(1 - r_{ik}^2 - r_{ij}^2 - r_{jk}^2)}{n}. \quad (14)$$

Once the asymptotic covariances and variances are available, Equation 13 can be applied to estimate the SE of the standardized indirect effect. When the SE s for the standardized indirect effect are available, Wald CIs may be constructed by Equation 5.

A structural equation modeling approach. An SEM approach may also be used to obtain the standardized indirect effect and its CI (see Figure 2). Using conventional SEM notations, squares and circles represent the observed variables and the latent variables, respectively. The double arrow on a variable represents the variance or error variance of that variable, depending on whether it is an independent or a dependent variable.

There are three important points that need to be noted here. First, the parameters f , g , and h in Figure 2 represent the standard deviations of the observed variables X , M , and Y , respectively. Second, there is no measurement error for the observed variables X , M , and Y . Third, the latent variance of Q is fixed at 1, whereas the following two nonlinear constraints are required to fix the variances of R and S constrained at 1:

$$d = 1 - a^2$$

and

$$e = 1 - b^2 - c'^2 - 2abc'. \quad (15)$$

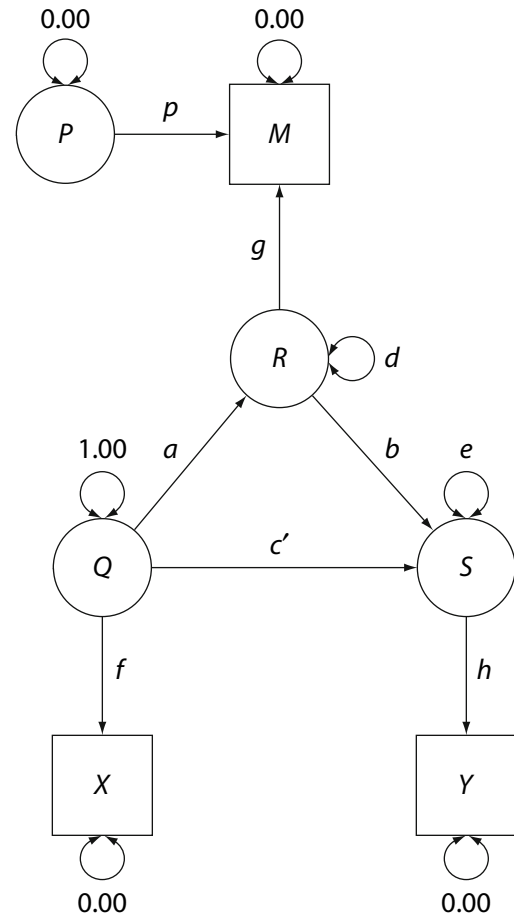


Figure 2. A structural equation model for the standardized indirect effect.

By using this parameterization, Q , R , and S become the standardized variables of X , M , and Y , respectively (Jöreskog & Sörbom, 1996). It should be noted that these constraints also ensure that the bootstrapped samples are standardized. Thus, a percentile bootstrap and BC bootstrap CIs of the standardized indirect effect can be obtained correctly by using this setup.

The standardized indirect effect can be easily estimated by the product term ab in Figure 2. To estimate the CI of the standardized indirect effect, we may introduce a phantom variable P (see, e.g., M. W.-L. Cheung, 2007, 2009; Raykov & Shrout, 2002; Rindskopf, 1984). The variance of P is fixed at 0. Thus, the addition of P has no effect on the parameter estimates and the model fit. It is introduced to directly estimate the product term of a and b ; that is,

$$p = ab. \quad (16)$$

The CI on p is equivalent to the estimated CI on the standardized indirect effect. Most SEM packages, such as LISREL, Mx (Neale, Boker, Xie, & Maes, 2005), and Mplus,

$$SE_{\text{Bobko-Rieck}} = \sqrt{s^2 \text{var}(r_{XM}) + t^2 \text{var}(r_{XY}) + u^2 \text{var}(r_{MY}) + 2st \text{cov}(r_{XM}, r_{XY}) + 2tu \text{cov}(r_{XY}, r_{MY}) + 2su \text{cov}(r_{XM}, r_{MY})} \quad (13)$$

allow for the creation of new parameters that are functions of other parameters (see Appendix B for the sample codes). Both LISREL (Jöreskog & Sörbom, 1996, pp. 345–347) and Mplus (Muthén & Muthén, 2007, pp. 553 and 577) use the multivariate delta method to calculate the *SEs* of the parameters with constraints. By using the SEM approach, a variety of methods for constructing the CIs of the standardized indirect effect are available to applied researchers.

An Illustration With an Empirical Example

To illustrate the procedures for the present article, an example used in M. W.-L. Cheung (2009) was adopted. In this example, job satisfaction was used as a mediator to explain the effect from job autonomy to life satisfaction. The descriptive statistics are shown in Table 1. Sample Mplus, LISREL, and Mx syntax using the SEM approach to estimate the standardized indirect effect is listed in Appendix B. The key point in obtaining a correct bootstrap CI on the standardized indirect effect is to standardize each bootstrapped sample, not the original data, before calculating the bootstrap CIs. According to Sobel's (1982) formula, the unstandardized indirect effect and its *SE* were 0.1232 and 0.0299, respectively. Thus, it is statistically significant ($z = 4.12, p < .001$). Since the *SDs* of the variables deviate from 1, interpreting the strength of the indirect effect is difficult.

According to Equation 6, the standardized indirect effect is 0.1991. According to M. W.-L. Cheung (2007; see note 2), the effect size can be qualified as small. The Wald CIs for Sobel-fixed, Aroian-fixed, Sobel-random, Aroian-random, Bobko-Rieck, and SEM-Wald are (0.1038, 0.2944), (0.1034, 0.2949), (0.0999, 0.2983), (0.0995, 0.2988), (0.0782, 0.3200), and (0.1085, 0.2897), respectively; whereas the naive bootstrap, the percentile bootstrap, the BC bootstrap, the likelihood-based, and the PRODCLIN CIs are (0.1059, 0.3106), (0.1089, 0.2967), (0.1113, 0.2985), (0.1117, 0.2942), and (0.1091, 0.2996), respectively. Although all CIs do not include 0, there appear to be moderate differences among these estimated values. However, judging which one is more accurate is difficult. The following simulation study addressed the accuracy of different CI estimators of the standardized indirect effect via a computer simulation.

A Simulation Study

Method

Multivariate normal data with known population standardized indirect effects were generated from the statistical environment R (R Development Core Team, 2008). Procedures for estimating the Wald

CIs and the naive bootstrap CI of the standardized indirect effects were implemented in R. A PRODCLIN program with the R interface (MacKinnon, Fritz, et al., 2007) was used to construct the PRODCLIN CI for the standardized data. The generated raw data were exported. Mplus was used to estimate the percentile bootstrap and the BC bootstrap CIs with the SEM approach. Since Mx is the only SEM package that provides a likelihood-based CI, it was used to obtain the likelihood-based CI of the standardized indirect effect. Five thousand replications per condition were used to ensure the stability of the findings. For the bootstrap CIs, 2,000 bootstrapped samples were used.

Effect sizes for the population standardized indirect effects.

Four levels of the standardized indirect effects were studied: zero, small, medium, and large. According to Cohen (1988, pp. 412–414; see also note 2), the levels of small (2% of the variance in the dependent variable), medium (13% of the variance in the dependent variable), and large (26% of the variance in the dependent variable) effect sizes of the standardized indirect effect were selected in the simulation studies. The population direct effect was fixed at zero in the simulation. The regression coefficients of *a* and *b* in Figure 1 were assumed to be the same. Since the population direct effect was 0 and all variables were standardized, the percentage of the variance explained for the dependent variable by the independent variable via the mediator was the square of the standardized indirect effect. Thus, the path coefficients for *a* and *b* that were assumed to be the same in the zero, small, medium, and large effect-size conditions were 0, 0.37, 0.60, and 0.71, respectively (see M. W.-L. Cheung, 2007; MacKinnon et al., 2002; MacKinnon et al., 2004, for similar designs).

Sample size. Four levels of sample sizes were studied: 50, 100, 200, and 500. These levels were selected to reflect common sample sizes used in typical research settings.

Population standard deviation of the variables. To assess whether the standard deviation of the variables would have any effect on performance, the raw data generated were multiplied by the corresponding population standard deviations. Two levels were studied: 1 and 5.

Methods of estimating CIs. Eleven methods for estimating the CIs of the standardized indirect effect were studied. Six of them were Wald CIs: Sobel-fixed, Aroian-fixed, Sobel-random, Aroian-random, Bobko-Rieck, and SEM-Wald. The other five methods were the naive bootstrap CI, the percentile bootstrap CI, the BC bootstrap CI, the likelihood-based CI, and the PRODCLIN CI.

Assessment of the empirical performance. The parameter estimates of the unstandardized indirect effect are generally unbiased (MacKinnon et al., 1995). Given that the standardized indirect effects are transformed directly from the unstandardized indirect effects, they are also unbiased. Since the model with one mediator is saturated, the parameter estimates based on the regression analysis and the SEM approach are the same. Therefore, the parameter estimates were not compared in the present study.

The empirical coverage probabilities of the constructed CIs were reported. If the constructed CIs are accurate, it can be expected that about 95% of the CIs that are constructed will include the population standardized indirect effect. A 95% CI is conservative if the empirical coverage probability is larger than 95%; it is liberal if the coverage probability is smaller than 95%. When the CI is applied to hypothesis testing, a conservative (and liberal) CI indicates that the Type I error rate is smaller (and larger) than the predefined level. Neither of them is desirable.

Results and Discussion

The empirical coverage probabilities of the 95% CIs of the standardized indirect effect are summarized in Table 2. Several observations can be made. First, the naive bootstrap CI is quite comparable to the percentile bootstrap and the BC bootstrap CI, except when the effect size is large. When the effect size is large, the naive bootstrap CI is too conservative. Another problem with the naive boot-

Table 1
Descriptive Statistics of the Variables

	LS	JS	JA
LS	1.0000		
JS	0.3704	1.0000	
JA	0.2148	0.5504	1.0000
<i>SD</i>	1.4868	2.0094	2.4034

Note—*N* = 200. LS, life satisfaction (dependent variable); JS, job satisfaction (mediator); JA, job autonomy (independent variable).

Table 2
Empirical Coverage Probabilities of 95% Confidence Intervals of the Standardized Indirect Effects

Standardized Indirect Effect	Sample Size	Sobel-Fixed	Aroian-Fixed	Sobel-Random	Aroian-Random	Bobko-Rieck	SEM-Wald	Naive Bootstrap	Percentile Bootstrap	BC Bootstrap	Likelihood-Based	PRODCLIN
Population $SD = 1$												
0.00	50	1.000	1.000	1.000	1.000	1.000	1.000	.997	.998	.991	.996	.997
	100	1.000	1.000	1.000	1.000	1.000	1.000	.998	.998	.994	.996	.996
	200	1.000	1.000	1.000	1.000	1.000	1.000	.998	.999	.995	.998	.998
	500	1.000	1.000	1.000	1.000	1.000	1.000	.998	.999	.997	.999	.998
0.14	50	.924	.936	.928	.939	.932	.915	.945	.940	.953	.945	.941
	100	.945	.949	.950	.953	.952	.936	.952	.947	.953	.948	.955
	200	.946	.949	.951	.954	.955	.936	.950	.944	.951	.946	.954
	500	.954	.956	.963	.963	.968	.946	.954	.949	.949	.951	.960
0.36	50	.958	.961	.972	.974	.976	.938	.959	.944	.949	.949	.966
	100	.961	.962	.974	.975	.981	.939	.964	.943	.945	.944	.969
	200	.965	.966	.978	.978	.986	.946	.966	.946	.946	.947	.969
	500	.971	.971	.982	.982	.993	.948	.967	.948	.948	.948	.968
0.51	50	.968	.969	.982	.982	.988	.938	.963	.937	.936	.942	.969
	100	.973	.975	.987	.987	.993	.949	.972	.947	.947	.949	.973
	200	.972	.972	.988	.988	.995	.946	.972	.945	.947	.946	.974
	500	.974	.974	.990	.990	.998	.947	.973	.945	.946	.947	.976
Population $SD = 5$												
0.00	50	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.994	.997	.998
	100	1.000	1.000	1.000	1.000	1.000	1.000	.998	.998	.994	.997	.998
	200	1.000	1.000	1.000	1.000	1.000	1.000	.998	.998	.996	.997	.998
	500	1.000	1.000	1.000	1.000	1.000	1.000	.999	.998	.993	.996	.997
0.14	50	.930	.944	.934	.946	.937	.920	.946	.941	.950	.948	.943
	100	.936	.942	.942	.947	.948	.929	.957	.942	.953	.948	.949
	200	.947	.950	.956	.958	.961	.939	.946	.947	.950	.950	.954
	500	.952	.952	.959	.959	.966	.943	.954	.947	.949	.947	.956
0.36	50	.954	.958	.966	.968	.971	.930	.959	.938	.945	.942	.965
	100	.963	.965	.976	.977	.984	.936	.964	.944	.947	.946	.968
	200	.968	.968	.982	.983	.990	.945	.966	.948	.950	.950	.972
	500	.973	.973	.986	.986	.993	.950	.974	.949	.950	.951	.973
0.51	50	.971	.971	.983	.984	.990	.944	.966	.946	.947	.948	.975
	100	.966	.967	.982	.982	.990	.936	.969	.938	.939	.938	.967
	200	.976	.976	.988	.988	.996	.952	.972	.950	.951	.953	.978
	500	.974	.974	.990	.990	.997	.947	.973	.948	.949	.948	.974

strap CI is that the bias is consistent across sample sizes. This means that using a larger sample size does not reduce the bias introduced in the naive bootstrap CI. Given that the naive bootstrap CI is not theoretically justified, it is not advisable to use it.

Second, the percentile bootstrap, the BC bootstrap, and the likelihood-based CIs had the best coverage probability when the population standardized indirect effect was larger than zero. Their empirical coverage is very close to 95%. This result is consistent with the general findings on the unstandardized indirect effect (see, e.g., M. W.-L. Cheung, 2007; MacKinnon et al., 2004).

Third, the empirical coverage probability of the PRODCLIN CI was slightly conservative when the population standardized indirect effect was medium or large. One speculation is that the current naive application of PRODCLIN CI does not take the standardization effect into account; it simply treats the standardized coefficients as unstandardized coefficients. It is well known that the sampling distributions between a covariance and a correlation are different. Further research may address how to improve the PRODCLIN CI for testing the standardized indirect effect.

Fourth, all Wald CIs were not as good as the bootstrap and the likelihood-based CIs. Among the Wald CIs, the

SEM-Wald CI was slightly better than the other Wald CIs, especially when the population standardized effect size was large. Since the SEM-Wald CI and other Wald CIs are all based on the multivariate delta method, one may wonder why the SEM-Wald CI is better than other Wald CIs. One speculation is that a full asymptotic covariance matrix of all parameters is used in constructing an SEM-Wald CI. This means that the estimates of a , b , V_X , and V_Y are allowed to be correlated in the SEM approach. In estimating Sobel-fixed and Aroian-fixed CIs, V_X and V_Y are assumed fixed, whereas ab , V_X , and V_Y are assumed uncorrelated in estimating Sobel-random and Aroian-random CIs.⁶ The availability of a full asymptotic covariance matrix of all parameters slightly improves the performance of the SEM-Wald CI.

Last—and most important—the coverage probability of the constructed CIs was too high (or the constructed CIs were too conservative), regardless of the estimation methods, when the population standardized indirect effect was zero. This finding is consistent with those of previous simulations of the unstandardized indirect effect (see, e.g., MacKinnon et al., 2002; MacKinnon et al., 2004), in which the estimated SEs were less accurate when the population indirect effect was zero. Moreover, using bootstrap

and likelihood-based CIs did not help much in this condition. Further research is definitely needed to improve the coverage probability in testing the standardized indirect effect when the population indirect effect is zero.

Of the factors studied, it was found that the population standard deviation and the sample size had little impact on the empirical coverage of the estimated CIs. The most crucial factors affecting the performance were the estimation method and the population effect size of the standardized indirect effect.

Conclusions and General Discussion

In the present article, it was argued that the standardized indirect effect is a more appropriate measure of effect size in a mediation analysis. Eleven methods of constructing the CIs of the standardized indirect effect were reviewed and compared via a computer simulation. It was found that the percentile bootstrap, the BC percentile bootstrap, and the likelihood-based CIs performed best in terms of coverage probability. Researchers are advised to report the CI of the standardized indirect effect using some of these methods.

The conventions suggested by Cohen (1988) have been directly adopted in the present article. Following the results of M. W.-L. Cheung (2007), values of 0, 0.14, 0.36, and 0.51 for the standardized indirect effect are qualified as having an effect size of “zero,” “small,” “medium,” and “large,” respectively. There are several issues worth discussing. As Cohen (1988) himself stressed, “[t]his is an operation fraught with many dangers: The definitions are arbitrary, such qualitative concepts as ‘large’ are sometimes understood as absolute, sometimes as relative; and thus they run a risk of being misunderstood” (p. 12). Even though there are many problems with these conventions, Cohen (1988) continued to defend his conventions, stating that they “will be found to be reasonable by reasonable people” (p. 13).

Even if one accepts Cohen’s (1988) conventions as reasonable (0%, 2%, 13%, and 26% of the variance explained), one does not necessarily have to agree on applying them to the indirect effect, as was done in the present article. MacKinnon and his colleagues (e.g., MacKinnon et al., 2002) applied Cohen’s (1988) conventions to the regression paths. Thus, the magnitudes of the effect sizes defined by M. W.-L. Cheung (2007) and MacKinnon and his colleagues are different by a square root (see note 2 for the details). Which one is more appropriate is an open question for further clarification. As was emphasized by Kline (2004), one should not consider these conventions as “rules”; otherwise, one would encounter the same problem as in null hypothesis significance testing. Moreover, the conventions “do not apply across all behavioral research areas” (Kline, 2004, p. 133).

Kirk (1996) suggested that one should consider the practical significance of the findings in one’s research area. That is, one should look at whether the effect is large enough to be of value in a practical sense. When a field is mature enough with sufficient empirical studies, Kline (2004) recommended using a meta-analysis to operationally define small versus large effect sizes. Reasonable guidelines on the standardized indirect effect in a par-

ticular field may then be suggested. Although the extent to which the conventions above are representative can be (and should be) arguable, it should be noted that such arbitrary conventions cannot be even suggested by using an unstandardized indirect effect. The present study serves as a starting point to define the magnitude of effect in a mediation analysis.

When research questions become increasingly complicated, one single mediator may not be sufficient to explain the psychological processes among the variables. Multiple specific mediators (MacKinnon, 2000; Preacher & Hayes, 2008a; Williams & MacKinnon, 2008) and intermediate mediators (Taylor, MacKinnon, & Tein, 2008) may be required. It may also be necessary to test moderated mediation or mediated moderation (see, e.g., Edwards & Lambert, 2007; Morgan-Lopez & MacKinnon, 2006; Preacher, Rucker, & Hayes, 2007). Sometimes, variables of interest may span more than one level (e.g., Bauer, Preacher, & Gil, 2006; Pituch, Whittaker, & Stapleton, 2005). Much research is required to develop reasonable measures of effect size to quantify the indirect effect in these settings. Many of these methods can be easily extended to standardized indirect effects discussed in the present study.

The present article has addressed only the methods of estimating the CIs. Since meta-analysis is widely used as a methodology to define values of effect size observed in applied settings, future studies may investigate the synthesis of standardized indirect effects. Both a fixed- and a random-effects meta-analysis (Hedges & Olkin, 1985; Hedges & Vevea, 1998; Hunter & Schmidt, 2004) may be applied to the standardized indirect effects. Instead of synthesizing estimates of the indirect (and direct) effects, an alternative approach is to use meta-analytic structural equation modeling (M. W.-L. Cheung & Chan, 2005). Correlation matrices (M. W.-L. Cheung & Chan, 2005) or covariance matrices (M. W.-L. Cheung & Chan, 2009) are compared and synthesized first. A mediating model is then fitted to the pooled correlation (or covariance) matrix. Further research may compare the advantages and disadvantages of synthesizing the indirect and direct effects versus synthesizing the correlation matrices.

Although standardized indirect effects are comparable across studies, the estimated values may still be attenuated by the presence of statistical artifacts, such as measurement errors (Hoyle & Kenny, 1999) and restrictions in range. Hunter and Schmidt (2004) identified 11 statistical artifacts that can be corrected in a meta-analysis. Future studies may address how to correct these statistical artifacts and to synthesize the standardized indirect effects in the context of mediation analysis.

AUTHOR NOTE

This research was supported by the Academic Research Fund Tier 1 (R-581-000-064-112) from the Ministry of Education, Singapore. I thank Maggie Chan, Ivan Ong, and Yiu-Fai Yung for their comments on earlier versions of this article. I am also grateful to three anonymous reviewers for their helpful comments. Correspondence concerning this article should be addressed to M. W.-L. Cheung, Department of Psychology, Faculty of Arts and Social Sciences, National University of Singapore, Block AS4, Level 2, 9 Arts Link, Singapore 117570 (e-mail: mikewlcheung@nus.edu.sg).

REFERENCES

- ALWIN, D. F., & HAUSER, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, **40**, 37-47.
- AMERICAN PSYCHOLOGICAL ASSOCIATION. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.
- AROIAN, L. A. (1944). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics*, **18**, 265-271.
- BARON, R. M., & KENNY, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology*, **51**, 1173-1182.
- BAUER, D. J., PREACHER, K. J., & GIL, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multi-level models: New procedures and recommendations. *Psychological Methods*, **11**, 142-163.
- BENTLER, P. M. (2007). Can scientifically useful hypotheses be tested with correlations? *American Psychologist*, **62**, 772-782.
- BENTLER, P. M., & LEE, S. Y. (1983). Covariance structures under polynomial constraints: Applications to correlation and alpha-type structural models. *Journal of Educational Statistics*, **8**, 207-222.
- BOBKO, P., & RIECK, A. (1980). Large sample estimators for standard errors of functions of correlation coefficients. *Applied Psychological Measurement*, **4**, 385-398.
- BOLLEN, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- BOLLEN, K. A., & STINE, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, **20**, 115-140.
- CHEUNG, G. W., & LAU, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods*, **11**, 296-325.
- CHEUNG, M. W.-L. (2007). Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling*, **14**, 227-246.
- CHEUNG, M. W.-L. (2009). Constructing approximate confidence intervals for parameters with structural equation models. *Structural Equation Modeling*, **16**, 267-294.
- CHEUNG, M. W.-L., & CHAN, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, **10**, 40-64.
- CHEUNG, M. W.-L., & CHAN, W. (2009). A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling*, **16**, 28-53.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- COHEN, J. (1994). The world is round ($p < .05$). *American Psychologist*, **49**, 997-1003.
- COHEN, J., COHEN, P., WEST, S. G., & AIKEN, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- CUDECK, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, **105**, 317-327.
- CUMMING, G., FIDLER, F., LEONARD, M., KALINOWSKI, P., CHRISTIANSEN, A., KLEINIG, A., ET AL. (2007). Statistical reform in psychology: Is anything changing? *Psychological Science*, **18**, 230-232.
- DAVISON, A. C., & HINKLEY, D. V. (1997). *Bootstrap methods and their application*. New York: Cambridge University Press.
- EDWARDS, J. R., & LAMBERT, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, **12**, 1-22.
- FRTZ, M. S., & MACKINNON, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, **18**, 233-239.
- GOODMAN, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, **55**, 708-713.
- HARLOW, L. L., MULAİK, S. A., & STEIGER, J. H. (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.
- HEDGES, L. V., & OLKIN, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- HEDGES, L. V., & VEVEA, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, **3**, 486-504.
- HOYLE, R. H., & KENNY, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 195-222). Thousand Oaks, CA: Sage.
- HUNTER, J. E., & HAMILTON, M. A. (2002). The advantages of using standardized scores in causal analysis. *Human Communication Research*, **28**, 552-561.
- HUNTER, J. E., & SCHMIDT, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- JAMES, L. R., & BRETT, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, **69**, 307-321.
- JAMES, L. R., MULAİK, S. A., & BRETT, J. M. (2006). A tale of two methods. *Organizational Research Methods*, **9**, 233-244.
- JÖRESKOG, K. G. (1999). *How large can a standardized coefficient be?* Lincolnwood, IL: Scientific Software International. Available at www.ssicentral.com.
- JÖRESKOG, K. G., & SÖRBOM, D. (1996). *LISREL 8: A user's reference guide*. Chicago: Scientific Software International.
- JÖRESKOG, K. G., SÖRBOM, D., DU TOIT, S., & DU TOIT, M. (1999). *LISREL 8: New Statistical Features*. Chicago: Scientific Software International.
- JUDD, C. M., & KENNY, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, **5**, 602-619.
- KIRK, R. E. (1996). Practical significance: A concept whose time has come. *Educational & Psychological Measurement*, **56**, 746-759.
- KLINE, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- MACKINNON, D. P. (2000). Contrasts in multiple mediator models. In J. S. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 141-160). Mahwah, NJ: Erlbaum.
- MACKINNON, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.
- MACKINNON, D. P., FAIRCHILD, A. J., & FRITZ, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, **58**, 593-614.
- MACKINNON, D. P., FRITZ, M. S., WILLIAMS, J., & LOCKWOOD, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODLIN. *Behavior Research Methods*, **39**, 384-389.
- MACKINNON, D. P., KRULL, J. L., & LOCKWOOD, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, **1**, 173-181.
- 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**, 83-104.
- MACKINNON, D. P., LOCKWOOD, C. M., & WILLIAMS, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, **39**, 99-128.
- MACKINNON, D. P., WARSİ, G., & DWYER, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, **30**, 41-62.
- MORGAN-LOPEZ, A. A., & MACKINNON, D. P. (2006). Demonstration and evaluation of a method for assessing mediated moderation. *Behavior Research Methods*, **38**, 77-87.
- MUTHÉN, L. K., & MUTHÉN, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles: Muthén & Muthén.
- NEALE, M. C., BOKER, S. M., XIE, G., & MAES, H. H. (2005). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Virginia Commonwealth University, Department of Psychiatry.
- NEALE, M. C., & MILLER, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics*, **27**, 113-120.
- OLKIN, I., & FINN, J. D. (1995). Correlation redux. *Psychological Bulletin*, **118**, 155-164.
- OLKIN, I., & SIOTANI, M. (1976). Asymptotic distribution of functions of a correlation matrix. In S. Iida (Ed.), *Essays in probability and statistics* (pp. 235-251). Tokyo: Shinko Tsusho.
- PITUCH, K. A., WHITTAKER, T. A., & STAPLETON, L. M. (2005). A comparison of methods to test for mediation in multisite experiments. *Multivariate Behavioral Research*, **40**, 1-23.
- 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**, 717-731.

- PREACHER, K. J., & HAYES, A. F. (2008a). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, **40**, 879-891.
- PREACHER, K. J., & HAYES, A. F. (2008b). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13-54). Thousand Oaks, CA: Sage.
- PREACHER, K. J., RUCKER, D. D., & HAYES, A. F. (2007). Addressing moderated mediation hypothesis: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, **42**, 185-227.
- QUINONES-VIDAL, E., LOPEZ-GARCIA, J. J., PENARANDA-ORTEGA, M., & TORTOSA-GIL, F. (2004). The nature of social and personality psychology as reflected in *JPSP*, 1965-2000. *Journal of Personality & Social Psychology*, **86**, 435-452.
- RAYKOV, T., BRENNAN, M., REINHARDT, J. P., & HOROWITZ, A. (2008). Comparison of mediated effects: A correlation structure modeling approach. *Structural Equation Modeling*, **15**, 603-626.
- RAYKOV, T., & SHROUT, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, **9**, 195-212.
- R DEVELOPMENT CORE TEAM (2008). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at www.r-project.org.
- RINDSKOPF, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika*, **49**, 37-47.
- SHROUT, P. E., & BOLGER, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, **7**, 422-445.
- SOBEL, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, **13**, 290-312.
- SOBEL, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structural models. *Sociological Methodology*, **16**, 159-186.
- STUART, A., & ORD, J. K. (1994). *Kendall's advanced theory of statistics: Vol. 1. Distribution theory* (6th ed.). New York: Oxford University Press.
- TAMHANE, A. C., & DUNLOP, D. D. (2000). *Statistics and data analysis: From elementary to intermediate*. Upper Saddle River, NJ: Prentice Hall.
- TAYLOR, A. B., MACKINNON, D. P., & TEIN, J. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, **11**, 241-269.
- WANSBEEK, T., & MEIJER, E. (2000). *Measurement error and latent variables in econometrics*. New York: Elsevier.
- WILKINSON, L., TASK FORCE ON STATISTICAL INFERENCE, AMERICAN PSYCHOLOGICAL ASSOCIATION, SCIENCE DIRECTORATE, WASHINGTON DC (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologists*, **54**, 594-604.
- WILLIAMS, J., & MACKINNON, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling*, **15**, 23-51.
- YUNG, Y.-F. (2008, July). *Testing and contrasting mediation or indirect effects in SEM: An analytic approach and its implementation*. Paper presented at the 73rd Annual Meeting of the Psychometric Society, Durham, NH.

NOTES

1. The generic terms *indirect effect* and *mediating effect* are used interchangeably in the present article.

2. It should be noted that values of the "small-," "medium-," and "large-" effect sizes used in the simulation studies conducted by M. W.-L. Cheung (2007) and MacKinnon and his colleagues (Fritz & MacKinnon, 2007; MacKinnon, Fritz, et al., 2007; MacKinnon et al., 2002; MacKinnon et al., 2004; Taylor et al., 2008) were different. All of them adopted Cohen's (1988) conventions in regression analysis; that is, the effect sizes were defined by percentages of variance of 2%, 13%, and 26% explained on the dependent variable. There are two major differences between the definitions of M. W.-L. Cheung (2007) and those of MacKinnon and his colleagues (e.g., MacKinnon et al., 2002). First, M. W.-L. Cheung (2007) defined the aforementioned percentages as the effect sizes for the standardized indirect effect. That is, the corresponding standardized indirect effects for the "small," "medium," and "large" effect sizes are 0.14, 0.36, and 0.51, respectively. MacKinnon and his colleagues used 0.14, 0.39, and 0.59 as the effect sizes for the regression coefficients from X to M and from M to Y . In other words, the effect sizes defined by MacKinnon and his colleagues are the squared values of those defined by M. W.-L. Cheung (2007). Second—and most important—M. W.-L. Cheung (2007) defined the effect sizes in terms of standardized scores; that is, all variables (X , M , and Y in Figure 1) were standardized, whereas in their simulation studies, MacKinnon and his colleagues used 1 for the variances of the independent variable and the error variances (X , e , and f in Figure 1) (MacKinnon, 2008, p. 98; MacKinnon et al., 2004, p. 106). In other words, the values of 0.14, 0.39, and 0.59 in MacKinnon's work cannot be interpreted as values for the standardized indirect effects.

3. Mplus Version 5 (Muthén & Muthén, 2007) provides a command OUTPUT: STANDARDIZED to obtain the *SE* of the standardized variables. In combination with the command MODEL INDIRECT, it is possible to estimate the CI of the standardized indirect effect. Empirical verification shows that this is equivalent to the SEM approach proposed in the present article. Mplus users may consider this option to simplify the specifications of the model. TCALIS in SAS 9.2 also implements procedures to obtain the *SE* and Wald CI of the standardized indirect effect (Yung, 2008). Since this option is not available in other SEM packages, such as LISREL and Mx, the approach proposed in the present article is still very useful to general SEM users.

4. Strictly speaking, there should be "hats" over a , b , and *SE* in the formulas, since these are estimated values. To simplify the presentation, the hats are dropped in the present article.

5. MacKinnon (2008, pp. 84-85) proposed an effect-size measure termed the *standardized effect*, which is defined as $I_{\text{unstand.}}/V_Y^{0.5}$. It is different from the standardized indirect effect discussed in the present article by the factor of $V_X^{0.5}$. MacKinnon (2008) suggested that the standardized effect is similar to the standardized mean difference in the context of an ANOVA.

6. As one anonymous reviewer suggested, it is easy for one to check the covariance (or correlation) among a , b , V_X , and V_Y by inspecting the asymptotic covariance (or correlation) matrix with the SEM approach. The covariance between a and b is 0 in the single-mediator model (Sobel, 1982). However, it is expected that the covariances between a (or b) and V_X (or V_Y) will be nonzero.

APPENDIX A

Derivations of the Standard Errors of the Standardized Indirect Effects

The standardized indirect effect ($I_{\text{stand.}}$) is defined as

$$\begin{aligned} I_{\text{stand.}} &= a \frac{V_X^{0.5}}{V_M^{0.5}} b \frac{V_M^{0.5}}{V_Y^{0.5}} \\ &= ab \frac{V_X^{0.5}}{V_Y^{0.5}}, \end{aligned} \quad (\text{A1})$$

where a and b are the unstandardized regression coefficients in Figure 1, and V_X , V_M , and V_Y are the variances of X , M , and Y , respectively. If we treat both V_X and V_Y as fixed and known (see, e.g., Bollen & Stine, 1990, p. 121; Cohen et al., 2003, p. 86), the estimated SE of the standardized indirect effect SE_{fixed} can be easily obtained by

$$SE_{\text{stand.}} = \frac{V_X^{0.5}}{V_Y^{0.5}} SE_{ab}, \quad (\text{A2})$$

where SE_{ab} is the estimated SE of the unstandardized indirect effect. By substituting the SE_{ab} on the basis of the work of Sobel (1982) and Aroian (1944), we may derive the approximate SE s for the standardized indirect effect:

$$SE_{\text{Sobel-fixed}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2} \quad (\text{A3})$$

and

$$SE_{\text{Aroian-fixed}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + SE_a^2 SE_b^2}. \quad (\text{A4})$$

A second approach to deriving the SE of the standardized indirect effect is to treat V_X and V_Y as random variables that have their own sampling variances. The asymptotic sampling variances of V_X and V_Y can be approximated by

$$\frac{2V_X^2}{n-1}$$

and

$$\frac{2V_Y^2}{n-1},$$

respectively, where n is the sample size (Tamhane & Dunlop, 2000, pp. 177–178). In deriving the sampling variance of $SE_{\text{stand.-random}}$, it is more convenient to treat ab as a single term. Assuming that ab , V_X , and V_Y are independent, $SE_{\text{stand.-random}}$ can be approximated by using the multivariate delta method (e.g., Stuart & Ord, 1994),

$$\begin{aligned} SE_{\text{stand.-random}} &= \sqrt{\left(\frac{\partial I}{\partial(ab)} \right)^2 \text{var}(ab) + \left(\frac{\partial I}{\partial V_X} \right)^2 \text{var}(V_X) + \left(\frac{\partial I}{\partial V_Y} \right)^2 \text{var}(V_Y)} \\ &= \sqrt{\left(V_X^{0.5} V_Y^{-0.5} \right)^2 SE_{ab}^2 + \left(0.5(ab) V_X^{-0.5} V_Y^{-0.5} \right)^2 \left(\frac{2V_X^2}{n-1} \right) + \left(-0.5(ab) V_X^{0.5} V_Y^{-1.5} \right)^2 \left(\frac{2V_Y^2}{n-1} \right)} \\ &= \sqrt{\left(SE_{ab}^2 + \frac{a^2 b^2}{n-1} \right) \frac{V_X}{V_Y}}, \end{aligned} \quad (\text{A5})$$

where

$$\frac{\partial I}{\partial(z)}$$

is the partial derivative of $I_{\text{stand.}}$ with respect to z with the other variables held constant. By comparing Equation A5 against Equation A2, we may also express Equation A5 as

$$SE_{\text{stand.-random}} = \sqrt{SE_{\text{stand.-fixed}}^2 + \left(\frac{a^2 b^2}{n-1} \right) \frac{V_X}{V_Y}}. \quad (\text{A6})$$

By substituting Sobel's (1982) and Aroian's (1944) estimates of SE_{ab} , we have the approximate SE s of the standardized indirect effect:

$$SE_{\text{Sobel-random}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + \left(\frac{a^2 b^2}{n-1} \right)} \quad (\text{A7})$$

and

$$SE_{\text{Aroian-random}} = \frac{V_X^{0.5}}{V_Y^{0.5}} \sqrt{a^2 SE_b^2 + b^2 SE_a^2 + SE_a^2 SE_b^2 + \left(\frac{a^2 b^2}{n-1} \right)}. \quad (\text{A8})$$

APPENDIX A (Continued)

Structural equation modeling may also be used to obtain the *SE* of the standardized indirect effect, as is shown in Figure 2. To obtain the standardized indirect effect, we have to fix the variances of *Q*, *R*, and *S* at 1. The variance of *Q* can be easily fixed to 1, whereas nonlinear constraints are required to fix the variances of *R* and *S*:

$$\begin{aligned} R &= aQ + \varepsilon_R \\ \text{var}(R) &= a^2 \text{var}(Q) + \text{var}(\varepsilon_R) \\ 1 &= a^2(1.00) + d. \end{aligned} \quad (\text{A9})$$

Thus, $d = 1 - a^2$ is required to ensure that the variance of *R* is constrained at 1.

$$\begin{aligned} S &= bR + c'Q + \varepsilon_S \\ S &= b(aQ) + \varepsilon_R + c'Q + \varepsilon_S \\ \text{var}(S) &= (ab + c')^2 \text{var}(Q) + b^2 \text{var}(\varepsilon_R) + \text{var}(\varepsilon_S) \\ 1 &= a^2 b^2 + 2abc' + c'^2 + b^2(1 - a^2) + e. \end{aligned} \quad (\text{A10})$$

Thus, $e = 1 - b^2 - c'^2 - 2abc'$ is required to fix the variance of *S* at 1.

APPENDIX B

Mplus, LISREL, and Mx Syntax for Constructing the Confidence Interval of the Standardized Indirect Effect

```

TITLE: Mplus code
DATA: FILE IS raw_data.dat;           ! Raw data is required by bootstrap
VARIABLE: NAMES X M Y;                ! X: Independent variable;
                                      ! M: Mediator; Y: Dependent variable
      USEVARIABLES ARE ALL;
!ANALYSIS: BOOTSTRAP=2000;             ! Activate bootstrap CI
MODEL:
    S BY Y*;                          ! Dependent variable, h in Figure 2
    R BY M*;                          ! Mediator, g in Figure 2
    Q BY X*;                          ! Independent variable, f in Figure 2

    S ON R (b)                        ! b in Figure 2
        Q (c);                       ! c in Figure 2

    R ON Q* (a);                      ! a in Figure 2
    R* (d);                          ! d in Figure 2
    S* (e);                          ! e in Figure 2

    Q@1.0;                           ! Fix the variance of Q

    ! No measurement error
    X@0.0;
    M@0.0;
    Y@0.0;
MODEL CONSTRAINT:
    NEW (p);
    d = 1 - a**2;
    e = 1 - b**2 - c**2 - 2*a*b*c;
    P = a*b;                          ! Standardized indirect effect
OUTPUT:
    CINTERVAL(symmetric);             ! Wald CI
    ! CINTERVAL(bootstrap);           ! Bootstrap CI
    ! CINTERVAL(BCbootstrap);         ! BC Bootstrap CI
-----
TI LISREL code to estimate SE
DA NI=3 AP=1                          ! Create one additional parameter
CM FI=raw_data.dat                    ! to store the standardized indirect effect
LA
X M Y
SE
M Y X /                               ! Arrange into Y-variables and X-variables
MO NY=2 NX=1 NE=2 NK=1 LX=DI,FR LY=DI,FR BE=SD PH=ID PS=DI TE=ZE TD=ZE

```

APPENDIX B (Continued)

```
! Constraints to standardize the mediator and the dependent variable
CO PS(1,1) = 1 - GA(1,1)**2
CO PS(2,2) = 1 - BE(2,1)**2 - GA(2,1)**2 - 2*GA(1,1)*BE(2,1)*GA(2,1)
```

```
! Standardized indirect effect
CO PAR(1) = GA(1,1)*BE(2,1)
OU
```

```
-----
Mx: Likelihood based CI on the standardized indirect effect
NGroups=2
Data NInput=3
RECTANGULAR FILE=raw_data.dat
LABELS X M Y
```

```
Begin Matrices;
```

```
  S Symm 6 6
```

```
  A Full 6 6
```

```
  F Full 3 6
```

```
  I Iden 6 6
```

```
  J Full 1 4
```

```
  V Iden 3 3
```

```
  K Full 1 4
```

```
  N Full 1 4
```

```
  M Full 6 1
```

```
End Matrices;
```

```
Matrix S
```

```
  0
```

```
  0 0
```

```
  0 0 0
```

```
  0 0 0 1
```

```
  0 0 0 0 0.2
```

```
  0 0 0 0 0 0.2
```

```
Specify S
```

```
  0
```

```
  0 0
```

```
  0 0 0
```

```
  0 0 0 0
```

```
  0 0 0 0 1
```

```
  0 0 0 0 0 2
```

```
Label Row S
```

```
  X M Y D F H
```

```
Label Col S
```

```
  X M Y D F H
```

```
Matrix A
```

```
  0 0 0 1 0 0
```

```
  0 0 0 0 1 0
```

```
  0 0 0 0 0 1
```

```
  0 0 0 0 0 0
```

```
  0 0 0 0.5 0 0
```

```
  0 0 0 0.5 0.5 0
```

```
Specify A
```

```
  0 0 0 4 0 0
```

```
  0 0 0 0 3 0
```

```
  0 0 0 0 0 5
```

```
  0 0 0 0 0 0
```

```
  0 0 0 6 0 0
```

```
  0 0 0 7 8 0
```

```
Label Row A
```

```
  X M Y D F H
```

```
Label Col A
```

```
  X M Y D F H
```

```

Matrix F
  1 0 0 0 0 0
  0 1 0 0 0 0
  0 0 1 0 0 0
Label Row F
  X M Y
Label Col F
  X M Y D F H

Matrix M
0
0
0
0
0
0
SPECIFY M
12
13
14
0
0
0

! Elements for constraints
Matrix J 4 4 6 6      ! elements for constraints
Matrix K 5 4 5 4      ! a: A(5,4)
Matrix N 6 5 6 5      ! b: A(6,5)

Begin Algebra;
  T=\part(A,J);
  U=\part(S,J);
  W=(V-T)~ & U;        ! standardized variables
  P=\part(A,K)*\part(A,N); ! standardized indirect effect
End Algebra;

MEANS F*M;
Covariance F & ((I - A)~ & S);
Interval P 1 1
End Group;

Constraint Group for the standardized indirect effect
Constraint
Begin Matrices = Group 1
  Z Unit 1 3
End Matrices;

Constraint \d2v(W)=Z;
End Group;

```
