Behav Res (2015) 47:484–493
DOI 10.3758/s13428-014-0485-8

# Multiple-group analysis approach to testing group difference in indirect effects

Ehri Ryu

Published online: 7 June 2014


**Abstract** This article introduces five methods that take a multiple-group analysis approach to testing a group difference in indirect effects. Unlike the general frameworks for testing moderated indirect effects, the five methods provide direct tests for equality of indirect effects between groups. A simulation study was conducted to examine the performance of the methods in terms of the empirical type I error rate, statistical power, and coverage of 95 % confidence intervals. The likelihood ratio test and percentile bootstrap confidence intervals are recommended. The methods are illustrated using an empirical data set.

**Keywords** Indirect effect · Moderated indirect effect · Multiple group analysis · Mediation

Mediation occurs when the relationship of an independent variable (X) to a dependent variable (Y) is transmitted through a mediator variable (M). Mediation analysis is widely used in social science research to investigate the process underlying the relationship between the independent and the dependent variables (Baron & Kenny, 1986; MacKinnon, 2008). Figure 1 depicts a simple mediation model that involves a single-mediator variable. In this model, the total effect of X on Y consists of two components: One is the effect of X on Y that is carried through M (i.e., the mediated effect), and the other is the effect of X on Y that is independent of M (i.e., the effect depicted by the single-headed arrow that directly links X and Y in Fig. 1). In the framework of path analysis, a mediated effect is referred to as an *indirect effect*—for example, the indirect effect of X on Y via M (Alwin & Hauser, 1975; Bollen, 1987; Fox, 1980).

E. Ryu (✉)
Department of Psychology, Boston College, Chestnut Hill, MA 02467, USA
e-mail: ehri.ryu.1@bc.edu

In mediation analysis, it is often a question of interest whether a mediated or an indirect effect is the same across different groups of individuals or under different conditions. In other words, it is a question whether the indirect effect is moderated by another variable (called a *moderator*). A special case of a moderated indirect effect is when the moderator is a categorical variable (e.g., gender, ethnicity group, age group). In this case, the question can be rephrased as comparing the indirect effect between groups or testing a group difference in the indirect effect.

The purpose of this article is to review statistical methods that take a multiple-group analysis approach to testing a group difference in indirect effects and to empirically evaluate their performance. Methods for testing indirect effects are reviewed, followed by a review of general frameworks that are commonly used for testing moderated indirect effects. Then the methods taking a multiple-group analysis approach are introduced. A simulation study evaluates the performance of the methods. The methods are illustrated using an empirical data set.

## Test of indirect effects

When the hypothesized model involves an indirect effect, research questions often require obtaining an estimate of the indirect effect and testing the significance of it. A number of statistical approaches have been proposed for testing an indirect effect. The causal steps procedure proposed by Kenny and his colleagues tests for the existence of an indirect effect by fulfilling a series of conditions (Baron & Kenny, 1986; Judd & Kenny, 1981; Kenny, Kashy, & Bolger, 1998). In MacKinnon, Lockwood, Hoffman, West, and Sheets's (2002) literature review, the causal steps procedure was the most commonly used approach in psychology. But the causal steps procedure has a limitation in that it provides only a binary conclusion as to

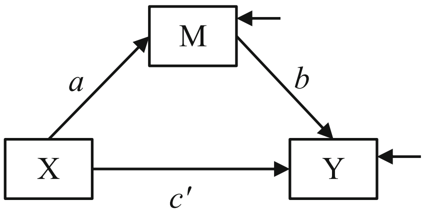


**Fig. 1** A single-mediator model

whether or not the indirect effect exists, rather than providing a formal test of the indirect effect. Also, the causal steps procedure has been shown to have low statistical power to detect the existence of an indirect effect (MacKinnon et al., 2002).

For the past decade, it has become a common practice in mediation analysis to formally test the indirect effect in addition to testing each of the individual paths that constitute the indirect effect. A formal test requires obtaining an estimate and standard error of confidence intervals (CIs) for the indirect effect. In the simple mediation model as shown in Fig. 1, the indirect effect of X on Y via M can be quantified as the product of $a$ and $b$.[1] In other words, an estimate of the indirect effect is obtained by $\widehat{ab}$ , which is referred to as the *product of coefficients* approach in the literature.

Various methods have been proposed and evaluated for obtaining the standard error or CIs for $ab$ (see MacKinnon et al., 2002; MacKinnon, Lockwood, & Williams, 2004; Preacher & Hayes, 2004, 2008). The first category of these methods is to obtain the standard error for $ab$ (Aroian, 1944; Sobel, 1982). Then the ratio of the estimate to the standard error is used in reference to the standard normal distribution to compute the significance of the indirect effect. The methods using standard errors assume that the sampling distribution of $ab$ is normal. But the assumption of a normal sampling distribution for the product is not valid even when the sampling distribution of each of the individual parameters $a$ and $b$ is normal (MacKinnon et al., 2002; MacKinnon et al., 2004). The distribution of $ab$ is skewed for nonzero indirect effect and the kurtosis also varies by the size of the indirect effect.

The second category of the methods is to obtain CIs for $ab$ that reflect the skewness (i.e., asymmetry) in the distribution of the product. Commonly used methods in this category are the *distribution of the product* method and the bootstrapping method. The distribution of the product method uses the table of the distribution of the product of two normal random variables by Meeker, Cornwell, and Aroian (1981) to produce asymmetric CIs for $ab$ (MacKinnon, Fritz, Williams, & Lockwood, 2007; MacKinnon et al., 2004). The asymmetric CIs have been shown to yield type I error rates closer to the nominal value than the traditional approach, which relies on the normality assumption.

The bootstrapping method takes a resampling approach to construct an empirical sampling distribution of $ab$ (Bollen & Stine, 1990; MacKinnon et al., 2004; Preacher & Hayes, 2004). A large number of bootstrap samples are drawn from the original sample with replacement. The size of the bootstrap samples is typically the same as the original sample size. For each bootstrap sample, the hypothesized model is estimated, and an estimate of the indirect effect of interest is collected. The collected set of estimates from a large number of bootstrap samples is used to construct a bootstrap sampling distribution. The CIs for $ab$ are obtained from the bootstrap sampling distribution. *Percentile bootstrap* $[100 * (1 - \alpha)]\%$ CIs are obtained by computing the $(\alpha/2)$ and $(1-\alpha/2)$ percentiles. *Bias-corrected bootstrap* CIs correct for bias in the central tendency of the estimate by adjusting the percentiles on the basis of the proportion of bootstrap samples producing $\widehat{ab}$ estimates lower than the original sample estimate.

In addition to the *distribution of the product* and the bootstrapping methods, a Monte Carlo method has been proposed to generate a sampling distribution of an indirect effect and to produce CIs (MacKinnon et al., 2004; Preacher & Selig, 2012). The Monte Carlo method assumes that the parameters $a$ and $b$ have a joint multivariate normal sampling distribution and generates values of $a$ and $b$ from a multivariate normal distribution. The estimates from the current sample are used as the parameter values for the joint multivariate normal distribution. The sampling distribution of $ab$ is constructed by generating $a$ and $b$ and computing their product for a large number of replications. The $(\alpha/2)$ and $(1-\alpha/2)$ percentiles provide $[100 * (1 - \alpha)]\%$ CIs. The Monte Carlo method is much less time consuming, as compared with the bootstrapping method, and it can be easily extended to a more complicated model, such as a model with multiple mediators or a multilevel mediation model. The Monte Carlo CIs for an indirect effect have been shown to perform as well as the bootstrapping methods (Preacher & Selig, 2012).

## Test of moderated indirect effects

Traditionally, there have been three approaches to testing moderated indirect effects (Edwards & Lambert, 2007).[2] In the *subgroup approach*, the sample is separated into subgroups based on the moderator variable, and the indirect effect

[1] The indirect effect of X on Y via M can also be quantified by the *difference in coefficients* ($c - c'$, where $c$ is the total effect of X on Y and $c'$ is the direct effect of X on Y). MacKinnon, Warsi, and Dwyer (1995) showed that the product of coefficients $ab$ and the difference in coefficients $c - c'$ yield identical estimates of the indirect effect when Y is continuous in a regression framework.

[2] Edwards and Lambert (2007) also reviewed the *piecemeal approach* as one of the general approaches to analyzing moderated indirect effects. The piecemeal approach is not reviewed in this article because it does not provide a test of moderated indirect effects. In the piecemeal approach, the indirect effect and moderated effect are analyzed separately, and the results are interpreted jointly.

is estimated in each group and compared between groups. The subgroup approach is appropriate when the moderator is a categorical variable. But for a general case, the subgroup approach has a limitation in that it requires an arbitrary categorization of a continuous moderator. The *moderated causal steps approach* follows the causal steps procedure of Baron and Kenny (1986) with the product term for the moderation effect (Muller, Judd, & Yzerbyt, 2005).

More recently, a number of studies have proposed general frameworks for analyzing moderated indirect effects (Edwards & Lambert, 2007; Fairchild & MacKinnon, 2009; MacKinnon, 2008; Preacher, Rucker, & Hayes, 2007). Although independently proposed, there is considerable overlap among these general frameworks. The general frameworks can analyze a moderation of a continuous variable on an indirect effect without requiring the continuous moderator to be arbitrarily categorized. The general frameworks can also handle a categorical moderator as long as it is appropriately coded.

In the general frameworks, a moderator variable and an appropriate product term are added in the model according to the specific hypothesis on which structural paths are moderated by the moderator. For example, in the single-mediator model, the indirect effect of X on Y via M is moderated by a moderator Z because path $a$ is conditional on the level of Z (depicted in Fig. 2), because path $b$ is conditional on the level of Z, or both. Then the conditional indirect effect given a level of Z can be obtained and tested. Bootstrapping methods and normal-theory standard errors have been proposed in the literature for testing the significance of a conditional indirect effect (Preacher et al., 2007). A simulation study by Preacher et al. (2007) showed that the bootstrapping methods yielded type I error rates closer to the nominal value and higher statistical power than the normal-theory standard error method. Although better than the normal-theory standard error method, the type I error rates of bootstrapping methods were still smaller than the nominal level (≤.011 in all conditions).

Note that the general frameworks provide a direct test of the moderation effect on a single structural path (e.g., moderation effect of Z on path $a$ is represented and tested by the structural path of X*Z on M—i.e., $a_3$ in Fig. 2) and a test of

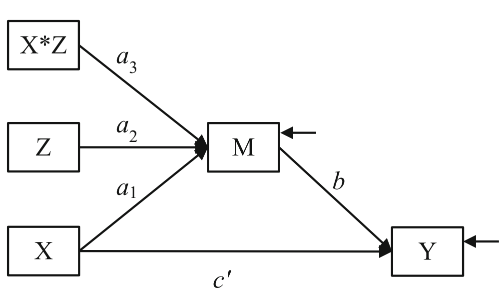


**Fig. 2** A moderated mediation model. Path $a$ is moderated by a moderator Z

the conditional indirect effect given a level of the moderator (e.g., indirect effect $ab$ at Z = $z$). But the general frameworks do not provide a direct test of the moderation of Z on the indirect effect $ab$. In the special case of a categorical moderator, the general frameworks can provide direct tests for whether or not the structural path $a$ is equal between groups and whether or not the indirect effect $ab$ is significantly different from zero in any particular group. But no direct test is provided for whether or not the indirect effect $ab$ is equal between groups.

## Multiple-group analysis approach to testing group difference in indirect effects

For the special case of a categorical moderator, a multiple-group analysis approach can be adopted to compare indirect effects between different groups. In a multiple-group analysis approach, the moderator no longer appears as a variable in the model. Instead, different levels of the moderator determine the group membership. Various methods can be used to test a group difference in indirect effects under a multiple-group analysis approach. For the ease of presentation, consider comparing the indirect effect $ab$ in the single-mediator model between two groups (denoted by G1 and G2, respectively), as shown in Fig. 3. The hypothesis can be written as $H_0$: $a_{(G1)}b_{(G1)}=a_{(G2)}b_{(G2)}$ or $H_0$:$a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}=0$.[3]

First, a likelihood ratio test (LR test; Bentler & Bonett, 1980; Bollen , 1989) can be used to test an equality constraint $a_{(G1)}b_{(G1)}=a_{(G2)}b_{(G2)}$. To obtain the LR test, two models need to be estimated: a model with no constraint ($M_0$) and another model with the equality constraint ($M_1$). The LR test between the two models results in a chi-square statistic with 1 degree of freedom ($df$ = 1).

$$\chi^2 = -2\log\left[\frac{L(M_1)}{L(M_0)}\right] = -2\log L(M_1) - [-2\log L(M_0)], \quad (1)$$

where $L(M_1)$ = likelihood of the model with the constraint $a_{(G1)}b_{(G1)}=a_{(G2)}b_{(G2)}$ and $L(M_0)$ = likelihood of the model with no constraint.

Alternatively, a Wald test can be used to test the equality constraint $a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}=0$ (Bollen, 1989; Wald, 1943). Unlike the LR test, only one model with no constraint

[3] Chan (2007) proposed a procedure for testing the hypothesis of equal indirect effect between groups without requiring a test of equality constraint. The procedure reparameterizes the model using a dummy latent variable so that the indirect effect is represented by a single parameter in the transformed model. Chan's procedure is not considered in the simulation study because direct tests of the equality constraint are available and, also, the primary focus of the present study is on the methods taking a multiple group analysis approach.

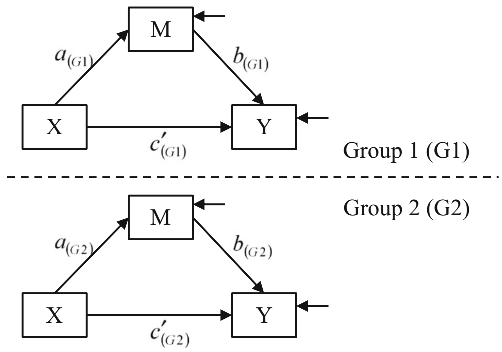


**Fig. 3** Multiple-group analysis for testing group difference in indirect effects between two groups

($M_0$) needs to be estimated in order to obtain the Wald statistic. The Wald statistic determines the extent to which the parameter estimates in the unconstrained model (e.g., $\widehat{a}_{(G1)}\widehat{b}_{(G1)}-\widehat{a}_{(G2)}\widehat{b}_{(G2)}$ in this case) differs from the constraint (e.g., zero in this case) with sampling error taken into account. The Wald statistic (W) is obtained by

$$W=\left[\boldsymbol{r}\left(\widehat{\boldsymbol{\theta}}_0\right)\right]'\left\{\left[\frac{\partial\boldsymbol{r}\left(\widehat{\boldsymbol{\theta}}_0\right)}{\partial\widehat{\boldsymbol{\theta}}_0}\right]'\left[\operatorname{acov}\left(\widehat{\boldsymbol{\theta}}_0\right)\right]\left[\frac{\partial\boldsymbol{r}\left(\widehat{\boldsymbol{\theta}}_0\right)}{\partial\widehat{\boldsymbol{\theta}}_0}\right]\right\}^{-1}\left[\boldsymbol{r}\left(\widehat{\boldsymbol{\theta}}_0\right)\right], \tag{2}$$

where $\boldsymbol{r}(\boldsymbol{\theta})$ = a vector of constraints, $\boldsymbol{r}\left(\widehat{\boldsymbol{\theta}}_0\right)$ = $\boldsymbol{r}(\boldsymbol{\theta})$ evaluated at $\widehat{\boldsymbol{\theta}}_0$ in which $\widehat{\boldsymbol{\theta}}_0$ = parameter estimates in the unconstrained model $M_0$, $\operatorname{acov}\left(\widehat{\boldsymbol{\theta}}_0\right)$ = estimated asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_0$ . The Wald statistic results in a chi-square statistic with $df$ = the number of constraints—that is, the dimension of $\boldsymbol{r}(\boldsymbol{\theta})$. When $\boldsymbol{r}(\boldsymbol{\theta})$ consists of a single constraint (e.g., $\theta_1 = 0$), the Wald statistic in Eq. 2 is simplified to

$$W=\frac{\widehat{\theta}_1^2}{\operatorname{avar}\left(\widehat{\theta}_1\right)}, \tag{3}$$

where $\operatorname{avar}\left(\widehat{\theta}_1\right)$ = estimated asymptotic variance of $\widehat{\theta}_1$ , and $df = 1$. For testing the difference in the indirect effect between G1 and G2, $\widehat{\theta}_1=\widehat{a}_{(G1)}\widehat{b}_{(G1)}-\widehat{a}_{(G2)}\widehat{b}_{(G2)}$ and $\operatorname{avar}\left(\widehat{\theta}_1\right)$ = estimated asymptotic variance of $\left[\widehat{a}_{(G1)}\widehat{b}_{(G1)}-\widehat{a}_{(G2)}\widehat{b}_{(G2)}\right]$ in the unconstrained model ($M_0$).

Second, bootstrapping methods can be used to compare indirect effects between groups. A large number (e.g., 1,000) of bootstrap samples are drawn from the original sample with replacement. In each bootstrap sample, the sample sizes of each group are the same as those in the original sample. In each bootstrap sample, estimates of $\widehat{a}^*_{(G1)}\widehat{b}^*_{(G1)}$ and $\widehat{a}^*_{(G2)}\widehat{b}^*_{(G2)}$ are obtained (* is used to indicate that the estimates are from bootstrap samples), and the group difference $\left[\widehat{a}^*_{(G1)}\widehat{b}^*_{(G1)}-\widehat{a}^*_{(G2)}\widehat{b}^*_{(G2)}\right]$ is calculated. The bootstrap sampling distribution for $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ is constructed using the set of values of $\left[\widehat{a}^*_{(G1)}\widehat{b}^*_{(G1)}-\widehat{a}^*_{(G2)}\widehat{b}^*_{(G2)}\right]$ collected from the bootstrap samples. Percentile bootstrap $[100 * (1-\alpha)]$% CIs for $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ are obtained by computing the ($\alpha/2$) and ($1-\alpha/2$) percentiles. Bias-corrected bootstrap CIs are obtained by using the adjusted percentiles $z'_{lower}=2z_0+z_{\alpha/2}$ and $z'_{upper}=2z_0+z_{1-\alpha/2}$, where $z_0$ is the proportion of bootstrap samples with $\left[\widehat{a}^*_{(G1)}\widehat{b}^*_{(G1)}-\widehat{a}^*_{(G2)}\widehat{b}^*_{(G2)}\right]$ smaller than the estimate of the group difference in the original sample $\left[\widehat{a}_{(G1)}\widehat{b}_{(G1)}-\widehat{a}_{(G2)}\widehat{b}_{(G2)}\right]$ .[4]

Third, a Monte Carlo method can be used to obtain CIs for $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$. It is assumed that the parameters $a_{(G1)}$, $b_{(G1)}$, $a_{(G2)}$, and $b_{(G2)}$ have a joint multivariate normal sampling distribution:

$$\begin{bmatrix} a_{(G1)} \\ b_{(G1)} \\ a_{(G2)} \\ b_{(G2)} \end{bmatrix}\sim MVN\left(\begin{bmatrix} \widehat{a}_{(G1)} \\ \widehat{b}_{(G1)} \\ \widehat{a}_{(G2)} \\ \widehat{b}_{(G2)} \end{bmatrix},\begin{bmatrix} \widehat{\sigma}^2_{a(G1)} & & & \\ \widehat{\sigma}_{\widehat{b}_{(G1)}\widehat{a}_{(G1)}} & \widehat{\sigma}^2_{b(G1)} & & \\ 0 & 0 & \widehat{\sigma}^2_{a(G2)} & \\ 0 & 0 & \widehat{\sigma}_{\widehat{b}_{(G2)}\widehat{a}_{(G2)}} & \widehat{\sigma}^2_{b(G2)} \end{bmatrix}\right). \tag{4}$$

The parameters in G1 are independent of the parameters in G2 because group 1 and group 2 are independent as long as the assumption of independent observations is valid. In the single-mediator model, the covariances between $a$ and $b$, $\widehat{\sigma}_{\widehat{b}_{(G1)}\widehat{a}_{(G1)}}$ and $\widehat{\sigma}_{\widehat{b}_{(G2)}\widehat{a}_{(G2)}}$ , are often replaced with zero (Preacher & Selig, 2012). As is shown in Eq. 4, the estimates and standard errors of $a$ and $b$ obtained in the original sample are used as parameter values in the multivariate normal distribution. For a large number of replications (e.g., 1,000), values of $a_{(G1)}$, $b_{(G1)}$, $a_{(G2)}$, and $b_{(G2)}$ are generated from the multivariate normal distribution shown in Eq. 4, and the group difference in the indirect effect $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ is calculated. An empirical sampling distribution of $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ is constructed using the generated values in a large number of replications. $[100 * (1-\alpha)]$% CIs for $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ are obtained by computing the ($\alpha/2$) and ($1-\alpha/2$) percentiles. Note that all of the methods introduced above provide direct tests for the group difference in the indirect effect.

[4] Bias-corrected and accelerated (BCa) bootstrap CIs can be obtained by considering the acceleration parameter in addition to bias correction. Previous simulation studies (e.g., Preacher et al., 2007; Preacher & Selig, 2012) showed that the difference between bias-corrected bootstrap CIs and BCa CIs was negligible in terms of type I error rates, statistical power, and coverage.

## Simulation

A simulation study was conducted to examine the performance of the methods in comparing an indirect effect between two groups. The model shown in Fig. 3 was used as the population model. Five different sets of parameter values were considered (see Table 1). In populations 1 and 2, there was no group difference in the indirect effect (*ab*). In populations 3, 4, and 5, the indirect effects were different between group 1 and group 2. The parameter value of $c' = 0.14$, the variance of X = 1, the residual variance of M = 1, and the residual variance of Y = 1 in both groups in all populations.[5] The data were generated assuming multivariate normality. All parameters in the mean structure are zero. Three sample sizes were considered: $N_{G1} = 100$ and $N_{G2} = 100$ (balanced), $N_{G1} = 150$ and $N_{G2} = 50$ (unbalanced 1), and $N_{G1} = 50$ and $N_{G2} = 150$ (unbalanced 2). With the unbalanced 1 sample size, the indirect effect is larger in the smaller group. With the unbalanced 2 sample size, the indirect effect is larger in the larger group. In each condition, 1,000 replications were conducted. Mplus 7 was used for data generation and model estimation (Muthén & Muthén, 1998–2012).

The following methods were used to test the group difference in the indirect effect: likelihood ratio test (LR) and Wald test (Wald) for equality constraint, percentile bootstrap 95 % CIs (PB), bias-corrected bootstrap 95 % CIs (BC), and Monte Carlo 95 % CIs (MC). The type I error rates (populations 1 and 2) and statistical power (populations 3, 4, and 5) were computed for each of the methods. For PB, BC, and MC, the coverage rates of the 95 % CIs were computed. For each method, the root mean square error (RMSE) of the coverage rates was computed across the five populations. The LR and Wald test statistics were obtained using Mplus 7. The PB, BC, and MC CIs were obtained using SAS 9.3.

For testing an indirect effect or a conditional indirect effect, it has been shown that the bias-corrected bootstrap method has a higher type I error rate and is more powerful than the percentile bootstrap method (MacKinnon et al., 2004; Preacher et al., 2007). Preacher and Selig (2012) showed that the Monte Carlo method performed as well as the percentile bootstrap and bias-corrected bootstrap methods in terms of coverage of CIs for an indirect effect. Note that the previous studies evaluated the performance of the methods in the context of testing an indirect effect (*ab*). The present simulation study evaluates the methods for testing a group difference in an indirect effect $[a_{(G1)}b_{(G1)} - a_{(G2)}b_{(G2)}]$.

**Table 1** Parameter values in population

| Population | Parameter | Group 1 | Group 2 |
|---|---|---|---|
| Population 1 | *a* | 0.39 | 0.39 |
| | *b* | 0.39 | 0.39 |
| | Proportion | .123 | .123 |
| Population 2 | *a* | 0.14 | 0.39 |
| | *b* | 0.39 | 0.14 |
| | Proportion | .046 | .052 |
| Population 3 | *a* | 0.14 | 0.39 |
| | *b* | 0.39 | 0.39 |
| | Proportion | .046 | .123 |
| Population 4 | *a* | 0.14 | 0.39 |
| | *b* | 0.14 | 0.39 |
| | Proportion | .019 | .123 |
| Population 5 | *a* | 0.14 | 0.59 |
| | *b* | 0.59 | 0.59 |
| | Proportion | .059 | .219 |

*Note*. See Fig. 1 for population model. $c' = 0.14$ in both groups in all populations; variance of X = 1; residual variance of M = 1; residual variance of Y = 1 in all populations. All parameters in the mean structure are zero. Proportion = proportion of variance in Y that is accounted for by the indirect effect *ab*

Table 2 shows the type I error rates and statistical power for testing a group difference in the indirect effect. The type I error rates smaller than .04 or greater than .06 are shown in bold, and 95 % CIs were computed for these type I error rates. [6] For both populations 1 and 2, the LR and PB methods yielded type I error rates close to the nominal value with all three sample sizes. The type I error rates of the Wald test were close to .05 in population 1 but smaller than .05 in population 2. In population 2, the 95 % CIs for type I error rates of the Wald test were [.025, .049] with the balanced sample size, [.018, .038] with the unbalanced 1 sample size, and [.022, .044] with the unbalanced 2 sample size. The type I error rates of the BC method were greater than .05 with the balanced sample size (95 % CI = [.047, .077] in population 1; 95 % CI = [.052, .082] in population 2). With the unbalanced 2 sample size, the type I error rates of the BC method was greater than .05 (95 % CI = [.050, .080]). The MC method yielded type I error rates close to the nominal value, except for population 1 with unbalanced sample sizes. With the unbalanced 1 sample size, the type I error rate was greater than .05 (95 % CI = [.046, .076]). With

[5] The population value of the direct effect ($c'$) was fixed at 0.14 in the present simulation study. Three additional sets of conditions were examined (not presented in this article): (1) $c' = 0.00$ with balanced sample size, (2) $c' = 0.00$ with unbalanced 1 sample size, and (3) $c' = 0.39$ with balanced sample size. The population value of the direct effect had little influence on the type I error rate, statistical power, and coverage rate of 95 % CIs for testing group difference in the indirect effect. The results for (1) and (3) were similar to those for the balanced sample size; the results for (2) were similar to those for the unbalanced 1 sample size in the present simulation study.

[6] The binomial 95 % CIs for type I error rate were computed by $\hat{p} \pm 1.96\sqrt{[\hat{p}(1-\hat{p})]/1000}$ (Agresti & Coull, 1998).

**Table 2** Type I error rates and power for testing group difference in indirect effect

| | LR | Wald | PB | BC | MC |
|---|---|---|---|---|---|
| Population 1 | | | | | |
| $N_{G1}$ = 100, $N_{G2}$ = 100 | .049 | .043 | .050 | **.062** | .047 |
| $N_{G1}$ = 150, $N_{G2}$ = 50 | .045 | .049 | .053 | .046 | **.061** |
| $N_{G1}$ = 50, $N_{G2}$ = 150 | .041 | .050 | .047 | .052 | **.023**[a] |
| Population 2 | | | | | |
| $N_{G1}$ = 100, $N_{G2}$ = 100 | .052 | **.037**[a] | .051 | **.067**[a] | .054 |
| $N_{G1}$ = 150, $N_{G2}$ = 50 | .044 | **.028**[a] | .041 | .057 | .052 |
| $N_{G1}$ = 50, $N_{G2}$ = 150 | .045 | **.033**[a] | .045 | **.065** | .040 |
| Population 3 | | | | | |
| $N_{G1}$ = 100, $N_{G2}$ = 100 | .288 | .268 | .286 | .306 | .169 |
| $N_{G1}$ = 150, $N_{G2}$ = 50 | .185 | .122 | .166 | .203 | .072 |
| $N_{G1}$ = 50, $N_{G2}$ = 150 | .246 | .300 | .252 | .268 | .170 |
| Population 4 | | | | | |
| $N_{G1}$ = 100, $N_{G2}$ = 100 | .648 | .601 | .621 | .665 | .522 |
| $N_{G1}$ = 150, $N_{G2}$ = 50 | .406 | .259 | .340 | .427 | .304 |
| $N_{G1}$ = 50, $N_{G2}$ = 150 | .575 | .662 | .580 | .587 | .436 |
| Population 5 | | | | | |
| $N_{G1}$ = 100, $N_{G2}$ = 100 | .750 | .738 | .732 | .741 | .578 |
| $N_{G1}$ = 150, $N_{G2}$ = 50 | .585 | .513 | .555 | .610 | .356 |
| $N_{G1}$ = 50, $N_{G2}$ = 150 | .601 | .644 | .619 | .593 | .520 |

*Note*. LR = likelihood ratio test for equality constraint; Wald = Wald test for equality constraint; PB = percentile bootstrap; BC = bias-corrected bootstrap; MC = Monte Carlo method. Type I error rates are shown for populations 1 and 2; power is shown for populations 3, 4, and 5. The type I error rates outside [.04, .06] are shown in bold

[a] The 95 % confidence interval for type I error rate does not include .05

the unbalanced 2 sample size, the type I error rate was smaller than .05 (95 % CI = [.014, .032]).[7]

Figure 4 shows the statistical power for testing a group difference in the indirect effect in populations 3, 4, and 5. Overall, the statistical power was higher with the balanced sample size than with unbalanced sample sizes. Between unbalanced sample sizes, the statistical power was higher with the unbalanced 2 sample size in which the indirect effect is larger in the larger group than with the unbalanced 1 sample size in which the indirect effect is larger in the smaller group. The MC method yielded the lowest empirical power among the five methods. The other four methods (LR, Wald, PB, and BC) yielded similar statistical power with the balanced sample size (Fig. 4a). With the unbalanced 1 sample size (Fig. 4b), the Wald test yielded a lower power than the LR, PB, and BC methods. With the unbalanced 2 sample size (Fig. 4c), the power of the Wald test was slightly higher than for the LR, PB, and BC methods. The BC method consistently showed a higher power than the PB, except for population 5 with the unbalanced 2 sample size. This advantage of the BC method was greater with the unbalanced 1 sample size in which the indirect effect was larger in the smaller group.

Table 3 shows the coverage rates of the 95 % CIs for the group difference in the indirect effect $[a_{(G1)}b_{(G1)}-a_{(G2)}b_{(G2)}]$ obtained by PB, BC, and MC methods. Note that for populations 1 and 2, in which there was no group difference in the indirect effect, the coverage rate is (1 − type I error rate). The coverage rates of the BC CIs were overall lower than .95 with the balanced sample size. This is the condition in which the BC method yielded inflated type I error rates. Both the inflated type I error rates and the lower coverage rates imply that the BC CIs were narrower than the desired width. The coverage rates of the MC CIs were slightly lower than .95 with the unbalanced 1 sample size. Again, this is the condition in which the type I error rates were higher than the nominal level. Table 3 also shows RMSE across all populations with each sample size. The RMSE values were small (≤.014) for all methods with all sample sizes.

In sum, the LR test performed well in all conditions with regard to type I error rate and power. The Wald test was less powerful than the LR test in all conditions when the sample sizes were equal in two groups (balanced) and when the sample size was larger in the group with smaller indirect effect (unbalanced 1). When the sample size was larger in the group with larger indirect effect (unbalanced 2), the Wald test was more powerful than the LR test. The type I error rate of the Wald test was smaller than the nominal level when the individual paths constituting the indirect effect were not equal between groups (i.e., population 2). The BC method was more powerful than the PB, except for one condition (population 5 with the unbalanced 2 sample size), but also showed inflated type I error rates. The MC method resulted in the lowest power among the five methods.

## Empirical example

Data from the PISA 2003 database (Programme for International Student Assessment, Organisation for Economic Co-operation and Development, 2004, 2005) were used to illustrate the methods for testing a group difference in indirect effects. The data were from 2,184 10th grade students (934 females, 1250 males) in Turkey.[8] The hypothesized model is shown in Fig. 5 (Yıldırım, 2012). The model examines the relationship of perceived teacher support with learning strategy use mediated by four motivational beliefs in mathematics.

[7] For population 1 with the unbalanced 2 sample size, the MC method was repeated for five additional times. The type I error rate of the MC method ranged from .038 to .048. The empirical type I error rate .023 in Table 2 occurred by chance.

[8] The multilevel structure of the data was ignored in the example. In this data set, 2,184 students were nested within 146 schools. The school size ranged from 2 to 35, mean school size = 14.96, and standard deviation = 8.171. The intraclass correlations were less than .077, except for math self-efficacy, for which the intraclass correlation was .165.

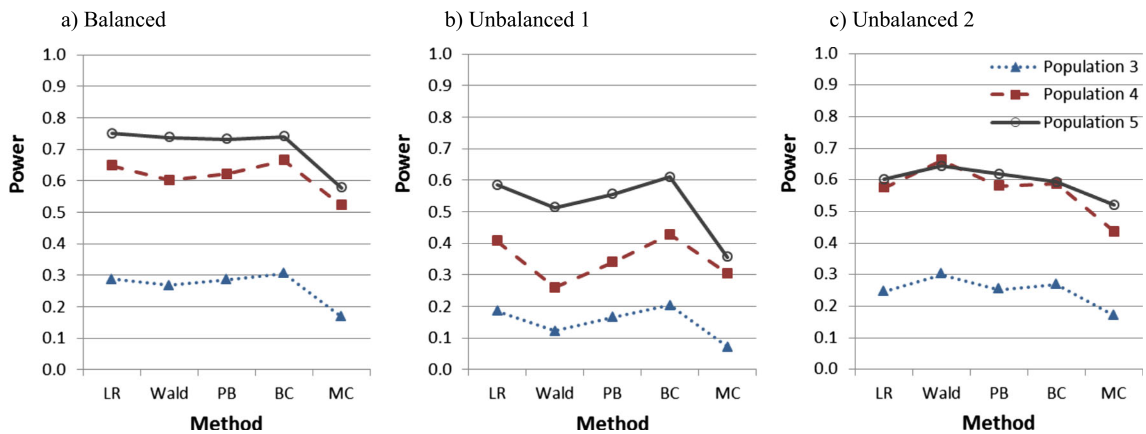


**Fig. 4** Statistical power for testing group difference in indirect effect in populations 3, 4, and 5. **a** Balanced: $N_{G1} = 100$, $N_{G2} = 100$. **b** Unbalanced 1: $N_{G1} = 150$, $N_{G2} = 50$. **c** Unbalanced 2: $N_{G1} = 50$, $N_{G2} = 150$. LR = likelihood ratio test for equality constraint; Wald = Wald test for equality constraint; PB = percentile bootstrap; BC = bias-corrected bootstrap; MC = Monte Carlo method

Perceived teacher support was measured by a composite score of items from a 4-point Likert-type scale in which a lower score indicated a *higher level of perceived teacher support* (Cronbach's alpha = .813). Each of the motivational beliefs was measured by a composite score of items from a 4-point Likert-type scale in which a lower score indicated a *higher level of motivational beliefs*: math self-efficacy (Cronbach's alpha = .813), anxiety (Cronbach's alpha = .813), intrinsic value (Cronbach's alpha = .896), and instrumental value (Cronbach's alpha = .842). Learning strategy use was measured by a composite score of items from a 4-point Likert-type scale in which a lower score indicated a *higher level of learning strategy use* (Cronbach's alpha = .860).

**Table 3** Coverage rates of 95 % confidence intervals

| | PB | BC | MC |
|---|---|---|---|
| $N_{G1} = 100$, $N_{G2} = 100$ | | | |
| Population 1 | .950 | .938 | .953 |
| Population 2 | .949 | .933 | .946 |
| Population 3 | .947 | .933 | .939 |
| Population 4 | .941 | .937 | .947 |
| Population 5 | .945 | .945 | .938 |
| RMSE | .005 | .014 | .008 |
| $N_{G1} = 150$, $N_{G2} = 50$ | | | |
| Population 1 | .947 | .954 | .939 |
| Population 2 | .959 | .943 | .948 |
| Population 3 | .946 | .950 | .936 |
| Population 4 | .948 | .950 | .944 |
| Population 5 | .941 | .948 | .939 |
| RMSE | .006 | .004 | .010 |
| $N_{G1} = 50$, $N_{G2} = 150$ | | | |
| Population 1 | .953 | .948 | .977 |
| Population 2 | .955 | .935 | .960 |
| Population 3 | .960 | .941 | .957 |
| Population 4 | .968 | .952 | .953 |
| Population 5 | .949 | .935 | .945 |
| RMSE | .010 | .010 | .014 |

*Note.* PB = percentile bootstrap; BC = bias-corrected bootstrap; MC = Monte Carlo method. RMSE = root mean square error across all populations

The model was just identified (df = 0). Table 4 shows the estimates and 95 % CIs for the indirect effects of perceived teacher support on learning strategy use via the four motivational beliefs in female and male students. The indirect effects were tested separately for each motivational beliefs variable. The 95 % CIs for each indirect effect were obtained using Sobel, percentile bootstrap, and bias-corrected bootstrap methods. For female students, the 95 % CIs did not include zero for all four indirect effects. For male students, the 95 % CIs did not include zero for the indirect effects via math self-efficacy, intrinsic value, and instrumental value. But for the indirect effect via anxiety, the 95 % CIs did include zero.

The indirect effects were compared between female and male students using the five methods: LR, Wald, PB, BC, and

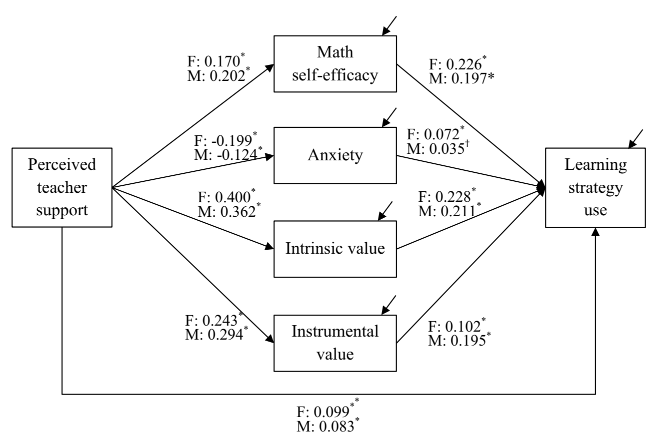


**Fig. 5** Path model of perceived teacher support, motivational beliefs, and learning strategy use (Yıldırım, 2012). The residuals for math self-efficacy, anxiety, intrinsic value, and instrumental value are correlated with one another. The estimates are shown. F = female; M = male. * $p < .05$; † $p = .062$

**Table 4** Estimates and 95 % confidence intervals for indirect effects

| Motivational Beliefs | Estimate | | | 95 % Confidence Intervals for Indirect Effect | | |
|---|---|---|---|---|---|---|
| | $\widehat{a}$ | $\widehat{b}$ | $\widehat{ab}$ | Sobel | Percentile Bootstrap | Bias-Corrected Bootstrap |
| Female ($N = 934$) | | | | | | |
| Math self-efficacy | 0.170 | 0.226 | 0.0384 | [0.023, 0.054] | [0.022, 0.057] | [0.023, 0.057] |
| Anxiety | −0.199 | 0.072 | −0.0143 | [−0.024, −0.004] | [−0.026, −0.005] | [−0.026, −0.006] |
| Intrinsic value | 0.400 | 0.228 | 0.0912 | [0.066, 0.116] | [0.065, 0.120] | [0.065, 0.123] |
| Instrumental value | 0.243 | 0.102 | 0.0248 | [0.012, 0.038] | [0.011, 0.043] | [0.010, 0.042] |
| Male ($N = 1250$) | | | | | | |
| Math self-efficacy | 0.202 | 0.197 | 0.0398 | [0.027, 0.052] | [0.025, 0.058] | [0.024, 0.058] |
| Anxiety | −0.124 | 0.035† | −0.0043 | [−0.009,0.001] | [−0.010, 0.001] | [−0.011, 0.001] |
| Intrinsic value | 0.362 | 0.211 | 0.0764 | [0.057, 0.096] | [0.056, 0.098] | [0.057, 0.100] |
| Instrumental value | 0.294 | 0.195 | 0.0573 | [0.041, 0.073] | [0.039, 0.077] | [0.039, 0.077] |

† $p = .062$; $p < .05$ for all the other estimates $\widehat{a}$ and $\widehat{b}$

MC (see Table 5). The LR and Wald tests concluded that the indirect effect via instrumental value was significantly smaller for female students than for male students. The indirect effects via the other three motivational beliefs were not significantly different at $p < .05$ between females and males. The PB, BC, and MC 95 % CIs did not include zero for the group difference in the indirect effect via instrumental value but did include zero for the other indirect effects. In sum, among the four motivational beliefs, only instrumental value showed a gender difference in the indirect effect of perceived teacher support on learning strategy use. Perceived teacher support has a stronger indirect effect on learning strategy use carried through instrumental value for male students.

## Discussion

This article examined the performance of the five methods that take a multiple-group analysis approach to testing group difference in indirect effects. Testing a group difference in indirect effects is a special case of testing moderated indirect effects in which the moderator is a categorical variable. The methods introduced in this article have an advantage in that they provide direct tests of a group difference in indirect effects that cannot be obtained in the general frameworks for testing moderated indirect effects.

Among the five methods, the LR and Wald tests have an advantage in that they do not require resampling of the data or Monte Carlo simulation to construct an empirical sampling distribution of the group difference in indirect effects. Another advantage of the LR and Wald tests is that they can be easily extended to more complicated situations. First, they can be used to test a group difference in indirect effects via multiple mediators such as X → M1 → M2 → Y. In this case, the indirect effect of X on Y through M1 and M2 consists of the three paths. The LR and Wald tests can be used to test $H_0$: $\beta_{M1,X(G1)}\beta_{M2,M1(G1)}\beta_{Y,M2(G1)}=\beta_{M1,X(G2)}\beta_{M2,M1(G2)}\beta_{Y,M2(G2)}$. Second, they can be used to simultaneously test a group difference in more than one indirect effect. For example, in the empirical example (Fig. 5), the LR and Wald tests can be

**Table 5** Difference in indirect effects between female and male students

| Motivational Beliefs | Difference | LR | Wald | 95 % Confidence Intervals for Difference in Indirect Effect | | |
|---|---|---|---|---|---|---|
| | | | | PB | BC | MC |
| Math self-efficacy | −0.0014 | 0.020 ($p = .889$) | 0.020 ($p = .888$) | [−0.025, 0.022] | [−0.025, 0.022] | [−0.023, 0.017] |
| Anxiety | −0.0100 | 3.502 ($p = .061$) | 3.146 ($p = .076$) | [−0.022, 0.001] | [−0.023, 0.001] | [−0.031, 0.013] |
| Intrinsic value | 0.0148 | 0.859 ($p = .354$) | 0.845 ($p = .359$) | [−0.019, 0.053] | [−0.019, 0.054] | [−0.001, 0.035] |
| Instrumental value | −0.0325 | 9.624 ($p = .002$) | 9.725 ($p = .002$) | [−0.059, −0.008] | [−0.059, −0.009] | [−0.051, −0.018] |

*Note*. Difference = difference in indirect effect between female and male groups ($\widehat{ab}$ in female − $\widehat{ab}$ in male). LR = likelihood ratio test for equality constraint; Wald = Wald test for equality constraint; PB = percentile bootstrap; BC = bias-corrected bootstrap; MC = Monte Carlo method. Chi-square statistics ($df = 1$) are shown for LR and Wald tests used to test an omnibus hypothesis $H_0$: $a_{k(G1)}b_{k(G1)}=a_{k(G2)}b_{k(G2)}$ for $k = 1, 2, 3$, and 4. Third, they can be used to compare indirect effects across more than two groups simultaneously. For example, the LR and Wald tests can be used to test a hypothesis $H_0$: $a_{(G1)}b_{(G1)}=a_{(G2)}b_{(G2)}=a_{(G3)}b_{(G3)}$.

In the simulation study, the LR test performed well in all conditions, but the Wald test showed a deflated type I error rate in some conditions. The Wald test statistic is obtained by a ratio of the squared estimate of the group difference in the indirect effect to the estimated asymptotic variance of the group difference in the indirect effect. The Wald test consistently yielded type I error rates lower than .05 in Population 2 with all three sample sizes. The lower type I error rates in population 2 imply that the asymptotic variance of the group difference in the indirect effect was overestimated when the individual paths constituting the indirect effect were different between groups. When the sample sizes were equal in both groups or when the sample size was larger in the group with smaller indirect effect, the LR test was more powerful. The Wald test was more powerful that the LR test when the sample size was smaller in the group with smaller indirect effect. Although the Wald test can be more powerful than the LR test under certain circumstances, the type I error rate of the Wald test may be deflated. The type I error rate of the LR test was close to the nominal level in all conditions. The LR test is recommended for testing the equality of indirect effects across groups.

The PB, BC, and MC methods provide CIs for the group difference in indirect effects. The use of CIs has been emphasized to convey the most complete meaning of empirical results (American Psychological Association, 2010; Wilkinson & the APA Task Force on Statistical Inference, 1999). The MC method has a practical advantage over the bootstrapping methods in that it is fast because it does not require resampling of the data and repeated model estimation. In a simulation study by Preacher and Selig (2012), for testing an indirect effect in a single group, the MC method performed equally well as, or better than, the other methods, including the percentile bootstrap and bias-corrected bootstrap, in terms of the coverage, CI width, and the symmetry of left-side and right-side misses. In the present study, the PB method overall performed the best in terms of coverage rates of the CIs. The coverage rates of the BC and MC CIs were not as close to .95 as the coverage rates of the PB method in some conditions. In terms of statistical power, the MC method was less powerful than the other methods in all conditions.

The PB method resulted in empirical type I error rates as close to or closer to the desired level than did the BC method. However, the BC method was more powerful than the PB method, and this advantage of the BC method was greater when the indirect effect was larger in the smaller group. Among the three methods producing CIs for a group difference in indirect effects, the bootstrapping methods are preferred to the MC method. The BC method is more powerful, but at the cost of inflated type I error rates. The PB method, although less powerful than BC, maintained the type I error rates at the desired level.

In conclusion, the performance of the LR method was comparable to or better than the performance of the other methods across all conditions. Considering the practical advantage and the quality of performance, the LR test is most recommended for testing the significance of a group difference in indirect effects. In addition, it is also recommended that the bootstrap (PB) CIs for the group difference in indirect effect are used along with the LR test.

## References


Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician, 52,* 119–126. doi:10.1080/00031305.1998.10480550

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40,* 35–47. doi:10.2307/2094445

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed., p. 33). 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. doi:10.1214/aoms/1177730442

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. doi:10.1037/0022-3514.51.6.1173

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis covariance structures. *Psychological Bulletin, 88,* 588–606.

Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. In C. C. Clogg (Ed.), *Sociological methodology 1987* (pp. 37–69). Washington, DC: American Sociological Association. doi:10.2307/271028

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. doi:10.2307/271084

Chan, W. (2007). Comparing indirect effects in SEM: A sequential model fitting method using covariance-equivalent specifications. *Structural Equation Modeling, 14,* 326–346. doi:10.1080/10705510709336749

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12,* 1–22. doi:10.1037/1082-989X.12.1.1

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science, 10,* 87–99. doi:10.1007/s11121-008-0109-6

Fox, J. (1980). Effect analysis in structural equation models. *Sociological Methods & Research, 9,* 3–26. doi:10.1177/004912418000900101

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

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology, In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233-265). New York: McGraw-Hill.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.

MacKinnon, D. P., Fritz, M. S., Williams, J. M., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods, 39,* 384–389. doi:10.3758/BF03193007

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediated and other intervening variable effects. *Psychological Methods, 7,* 83–104. doi:10.1037/1082-989X.7.1.83

MacKinnon, D. P., Lockwood, C. M., & Williams, J. M. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39,* 99–128. doi:10.1207/s15327906mbr3901_4

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

Meeker, W. Q., Cornwell, L. W., & Aroian, L. A. (1981). *Selected tables in mathematical statistics, vol. VII: The product of two normally distributed random variables*. Providence, RI: American Mathematical Society.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediatd and mediation is moderated. *Journal of Personality and Social Psychology, 89,* 852–863. doi:10.1037/0022-3514.89.6.852

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: Author.

Organisation for Economic Co-operation and Development. (2005). *PISA 2003 technical report*. Paris: Author.

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. doi:10.3758/BF03206553

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891. doi:10.3758/BRM.40.3.879

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185–227. doi:10.1080/00273170701341316

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6,* 77–98. doi:10.1080/19312458.2012.679848

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology 1982* (pp. 290–312). Washington, DC: American Sociological Association. doi:10.2307/270723

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society, 54,* 426–482. doi:10.2307/1990256

Wilkinson, L., & the APA Task Force on Statistical Inference, (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594–604. doi:10.1037/0003-066X.54.8.594

Yıldırım, S. (2012). Teacher support, motivation, learning strategy use, and achievement: A multilevel mediation model. *The Journal of Experimental Education, 80,* 150–172. doi:10.1080/00220973.2011.596855