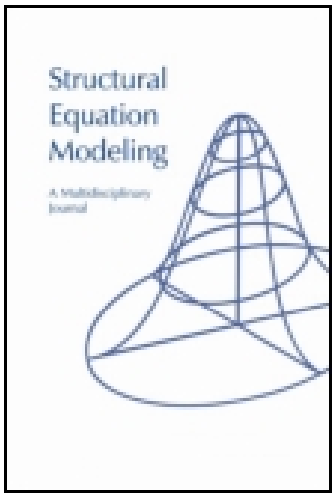

Structural Equation Modeling: A Multidisciplinary Journal

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hsem20

Causal Effects in Mediation Modeling: An Introduction With Applications to Latent Variables

Bengt Muthén[a] & Tihomir Asparouhov[a]

[a] Muthén & Muthén, Los Angeles, CA

Published online: 08 Oct 2014.

To cite this article: Bengt Muthén & Tihomir Asparouhov (2015) Causal Effects in Mediation Modeling: An Introduction With Applications to Latent Variables, Structural Equation Modeling: A Multidisciplinary Journal, 22:1, 12-23, DOI: 10.1080/10705511.2014.935843

To link to this article: http://dx.doi.org/10.1080/10705511.2014.935843

Structural Equation Modeling: A Multidisciplinary Journal, 22: 12–23, 2015

ISSN: 1070-5511 print / 1532-8007 online
DOI: 10.1080/10705511.2014.935843

Routledge
Taylor & Francis Group

Causal Effects in Mediation Modeling: An Introduction With Applications to Latent Variables

Bengt Muthén and Tihomir Asparouhov
Muthén & Muthén, Los Angeles, CA

Causal inference in mediation analysis offers counterfactually based causal definitions of direct and indirect effects, drawing on research by Robins, Greenland, Pearl, VanderWeele, Vansteelandt, Imai, and others. This type of mediation effect estimation is little known and seldom used among analysts using structural equation modeling (SEM). The aim of this article is to describe the new analysis opportunities in a way that is accessible to SEM analysts and show examples of how to perform the analyses. An application is presented with an extension to a latent mediator measured with multiple indicators.

Keywords: counterfactuals, randomization, structural equation modeling

Causal inference in mediation analysis offers counterfactually based causal definitions of direct and indirect effects, drawing on research by Robins, Greenland, Pearl, VanderWeele, Vansteelandt, Imai, and others. For overviews and references, see, for example, VanderWeele and Vansteelandt (2009), Valeri and VanderWeele (2013), and Muthén (2011).

To date, this type of mediation effect estimation is little known and seldom used among analysts using structural equation modeling (SEM). Part of the reason is that the literature is difficult to penetrate when coming from an SEM background. Causal inference in mediation analysis is often discussed by presenting the definitions of the causal effects using counterfactual and potential outcome concepts, typically not familiar to researchers using SEM. The aim of this article is to describe the new analysis opportunities in a way that is more accessible to SEM analysts and show examples of how to perform the analyses.

The next section discusses mediation effects in an intuitive way to provide the background for causal effect definitions, which are presented in the following section. The article then discusses the use of a latent variable mediator and presents a simulation study as well as an application with moderated mediation, before the article concludes. All analyses are carried out using M*plus*. The scripts are given in the Appendix. Several groups have recently produced causal inference software (see, e.g., Tingley, Yamamoto, Hirose, Keele, & Imai, 2013; Valeri & VanderWeele, 2013), but a unique feature of M*plus* is the ability to estimate the causal effects in the presence of latent variables.

Correspondence should be addressed to Bengt Muthén, Muthén & Muthén, 3463 Stoner Avenue, Los Angeles, CA 90066. E-mail: bmuthen@statmodel.com

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hsem.

THE ISSUES, INTUITIVELY

The effects presented in the causal inference literature in most cases coincide with effects that have traditionally been used in mediation analysis in the SEM tradition (MacKinnon, 2008). Important exceptions include mediation with a binary Y, a count Y, a binary M, and models with a treatment–mediator interaction. Before introducing the causal effect definitions using counterfactuals, the article first provides an intuitive description of the issues at hand using three specific cases: treatment–mediator interaction, binary Y, and count Y.

Continuous M and Y With Treatment–Mediator Interaction

Consider Figure 1, which corresponds to a randomized trial with a binary treatment dummy variable x (0 = control, 1 = treatment), a continuous mediator m, and a continuous outcome y, a situation examined in detail by MacKinnon

(2008). A special feature is that the treatment and mediator interact in their influence on the outcome y. This possibility is important to the so-called MacArthur approach to mediation (Kraemer, Kiernan, Essex, & Kupfer, 2008). As pointed out in, for example, VanderWeele and Vansteelandt (2009), the possibility of this interaction was emphasized in Judd and Kenny (1981) but not in the influential Baron and Kenny (1986) article on mediation, and is therefore often not explored. The interaction possibility is, however, stated in James and Brett (1984) and more recently in Preacher, Rucker, and Hayes (2007). In our experience the treatment–mediator effect is not often found in behavioral science applications, but the model provides a pedagogical way to introduce different mediation effects. The model of Figure 1 is used to first discuss the SEM concepts of direct and indirect effects and in a later section the corresponding causal concepts.

Assuming linear relationships, Figure 1 translates into

$$y_i = \beta_0 + \beta_1\ m_i + \beta_2\ x_i + \beta_3\, x_i\, m_i + \epsilon_{1i}, \tag{1}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \epsilon_{2i}, \tag{2}$$

where the residuals ϵ_1 and ϵ_2 are assumed normally distributed with zero means, variances σ_1^2, σ_2^2 and uncorrelated with each other and with the predictors in their equations. SEM considers the reduced form of this model, obtained by inserting Equation 2 in Equation 1,

$$\begin{aligned} y_i &= \beta_0 + \beta_1\,(\gamma_0 + \gamma_1\, x_i + \epsilon_{2i}) + \beta_2\ x_i \\ &\quad + \beta_3\, x_i\ (\gamma_0 + \gamma_1\, x_i + \epsilon_{2i}) + \epsilon_{1i}, \\ &= \beta_0 + \beta_1\ \gamma_0 + \beta_1\ \ \gamma_1\ x_i + \beta_3\ \gamma_0\ x_i + \beta_3\ \gamma_1\ x_i^2 \\ &\quad + \beta_2\, x_i + \beta_1\, \epsilon_{2i} + \beta_3\, x_i\, \epsilon_{2i} + \epsilon_{1i}, \end{aligned} \tag{3}$$

so that the expected value of y conditional on x is

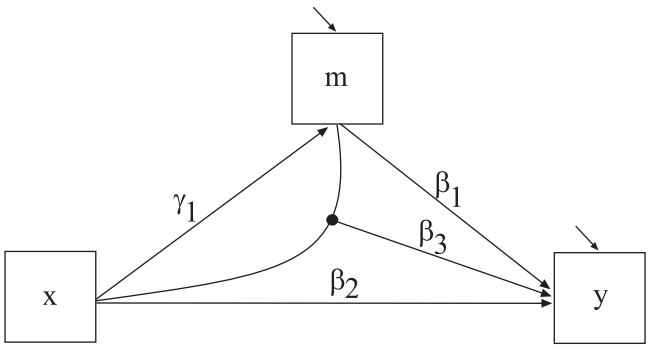

FIGURE 1 A mediation model with treatment–mediator interaction. The filled circle represents an interaction term consisting of the variables connected to it without arrow heads, in this case x and m.

$$E(y|x) = \beta_0 + \beta_1\ \gamma_0 + \beta_1\, \gamma_1\, x + \beta_3\, \gamma_0\ x + \beta_3\, \gamma_1\, x^2 + \beta_2\ x. \tag{4}$$

First, assume no treatment–mediator interaction; that is, $\beta_3 = 0$. In this case, considering the total effect difference $E(y|x=1) - E(y|x=0)$, one might deduce from Figure 1 and the reduced-form expression of Equation 4 that the indirect effect of x on y via m is $\beta_1\ \gamma_1$ and the direct effect is β_2. These are the standard formulas used in mediation modeling.

Second, let $\beta_3 \neq 0$, allowing treatment–mediator interaction. In this case, it is perhaps less clear how to deduce the direct and indirect effect. Using a term that will be defined later, one could consider the indirect effect to be

$$\textit{Total Natural Indirect Effect}\ (\textit{TNIE}) \ = \ \beta_1\ \gamma_1 + \beta_3\ \gamma_1, \tag{5}$$

a sum composed of a main part $\beta_1\ \gamma_1$ and an interaction part $\beta_3\ \gamma_1$. In this way, there can be a indirect effect even if $\beta_1 = 0$.

One could consider the direct effect to be

$$\textit{Pure Natural Direct Effect}\ (\textit{PNDE}) \ = \ \beta_2 + \beta_3\ \gamma_0, \tag{6}$$

where the second term is included because the γ_0 intercept is not part of the influence of x on m and is therefore not seen as part of the indirect effect. In this way, because of the treatment–mediator interaction there can be a direct effect even if $\beta_2 = 0$.

The sum of the direct and indirect effects is the total effect,

$$\textit{Total Effect}\ (\textit{TE}) \ = \ \textit{PNDE} + \textit{TNIE}. \tag{7}$$

These effects make intuitive sense and will later on be expressed in terms of causal effects derived using counterfactuals. At that point an alternative set of direct and indirect effects will also be presented using the indirect effect

$$\textit{Pure Natural Indirect Effect}\ (\textit{PNIE}) \ = \ \beta_1\ \gamma_1, \tag{8}$$

and the direct effect

$$\textit{Total Natural Direct Effect}\ (\textit{TNDE}) \ = \ \beta_2 + \beta_3\ \gamma_0 + \beta_3\ \gamma_1 \tag{9}$$

which also sum to the total effect,

$$\textit{Total Effect}\ (\textit{TE}) \ = \ \textit{TNDE} \ + \ \textit{PNIE}. \tag{10}$$

This second effect decomposition might seem less natural. One way to understand the two alternative decompositions is by starting from the indirect effect definitions. PNIE in Equation 8 does not include the interaction term so the effect

is pure in the sense of being due to the mediator alone. Not including the interaction, the interaction instead ends up in TNDE of Equation 9 to make the sum of the two effects add to the total effect (TE). The two different decompositions will be clarified when considering the counterfactual definitions. The decomposition using the effects TNIE of Equation 5 and PNDE of Equation 6 is the one most often considered in the literature and is the focus in the examples of this article.

Continuous *M* and Binary *Y*

As the next step in getting an intuitive grasp of mediation model effects, consider the case of a binary Y. As a starting point, the top part of Figure 2 shows a simple mediation model with a continuous mediator and final outcome. Because both the mediator and the final outcome are continuous, the indirect effect is $a \times b$. In the bottom part of Figure 2, y is instead a binary outcome and y^* is a continuous latent response variable behind y such that when y^* exceeds a threshold, $y = 1$ is observed rather than $y = 0$. The relationship between y and m is a probit or logit regression. The latent response variable formulation gives the same probit or logit regression for y related to m, but does so using a linear regression of y^* on m. A normally distributed residual in this regression results in probit and a logistically distributed residual results in logit. The residual variance is implicitly fixed at one with probit and at $\pi^2/3$ with logit. The coefficient b is a probit or a logit coefficient when considering y to be the dependent variable and a linear regression coefficient when considering y^* to be the dependent variable.

With a binary y as in the bottom part of Figure 2, the conventional $a \times b$ product formula for an indirect effect is only valid for the underlying continuous latent response variable y^*, not for the observed categorical outcome y itself (similarly, with a binary mediator, conventional product formulas for indirect effects are only valid for a continuous latent response variable behind the mediator). Earlier literature not referring to causal effects discussed indirect effects with y^* as the dependent variable (see, e.g., MacKinnon & Dwyer, 1993; MacKinnon, Lockwood, Brown, Wang, & Hoffman, 2007). It is clear that the product formula is valid for y^* because this case is the same as that of the top part of Figure 2; the dependent variable is continuous and we are considering linear regression. The causal effect literature, however, considers effects with y as the dependent variable. Why the product formula is not valid for y can be understood as follows.

The short answer to why there is a problem with using $a \times b$ as an indirect effect for a binary outcome is that of using only two parameters when six are needed. Consider Figure 3 and changes in the probability $P(y = 1|x)$. A unit change in x around the value x_1 changes the probability relatively little compared to a unit change in x around the value x_2 because the probability curve is steeper at the higher x value. In other words, a given change in x has a different impact on the probability depending on where on the probability curve the change takes place. This cannot be captured using only the two parameters of a and b; that is, the $a \times b$ indirect effect ignores the problem of nonconstant effects. This problem is avoided by focusing directly on the probability $P(y = 1|x)$, which as will be seen, is what is done with the causal indirect effect, using six parameters.

To help think about indirect effects in terms of $P(y = 1|x)$ intuitively, consider the following mediation model with a binary X, continuous M, binary Y, and for simplicity no treatment–mediator interaction:

$$y_i^* = \beta_0 + \beta_1\ m_i + \beta_2\ x_i + \epsilon_{1i}, \tag{11}$$

$$m_i = \gamma_0 + \gamma_1\ x_i + \epsilon_{2i}, \tag{12}$$

where the variance of ϵ_1 is fixed at c, c is 1 for probit and $\pi^2/3$ for logit, and the variance of ϵ_2 is denoted σ_2^2. Inserting Equation 12 into Equation 11 gives the reduced-form expectation and variance expressions

$$E(y^*|x) = \beta_0 + \beta_1\ \gamma_0 + \beta_1\ \gamma_1\ x + \beta_2 x, \tag{13}$$

$$V(y^*|x) = V(\epsilon_1 + \beta_1\ \epsilon_2) = c + \beta_1^2\ \sigma_2^2. \tag{14}$$

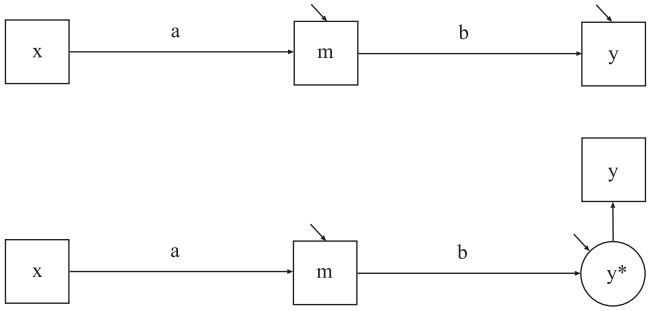

FIGURE 2 Indirect effect with continuous versus binary Y.

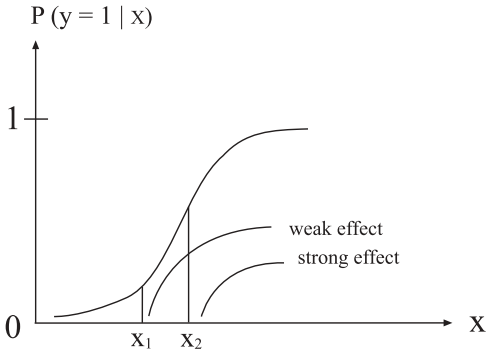

FIGURE 3 Nonconstant effects.

In the probit case, y^* conditional on x has a normal distribution because it involves the sum of two normal residuals ϵ_1 and ϵ_2. The probability can therefore be obtained using the standard normal distribution function Φ,

$$P(y = 1|x) = P(y^* > 0|x) = \Phi[E(y^*|x)/\sqrt{V(y^*|x)}]. \quad (15)$$

The six parameters mentioned earlier appear in this probability: $\beta_0, \beta_1, \gamma_0, \gamma_1, \beta_2, \sigma_2^2$. In the logit case, y^* conditional on x has a distribution that is a combination of a normal and a logistic residual and the probability can therefore not be expressed as succinctly as for probit. The probability can, however, be easily computed using integration over the normally distributed ϵ_2.

It is natural to think of effects for a binary outcome in terms of differences between probabilities for the treatment and control groups, $x = 1$ versus $x = 0$. The question is, however, how these differences should be expressed. It seems straightforward that the total effect is the difference in probabilities for $x = 1$ versus $x = 0$. The x values only come into play in $E(y^*|x)$, so it is instructive to focus on this part:

$$x = 1 : \beta_0 + \beta_1\ \gamma_0 + \beta_1\ \gamma_1 + \beta_2, \quad (16)$$

$$x = 0 : \beta_0 + \beta_1\ \gamma_0, \quad (17)$$

The $\beta_1\ \gamma_1 + \beta_2$ terms are what create the total effect when Equations 16 and 17 are inserted into Equation 15 for $x = 1$ compared to $x = 0$. Deciding on how to define the indirect effect is less clear. Should the difference be a function of only the product $\beta_1\ \gamma_1$ as when comparing

$$x = 1 : \beta_0 + \beta_1\,\gamma_0 + \beta_1\,\gamma_1, \quad (18)$$

$$x = 0 : \beta_0 + \beta_1\,\gamma_0, \quad (19)$$

or should the difference be a function also of β_2 as when comparing

$$x = 1 : \beta_0 + \beta_1\,\gamma_0 + \beta_2 + \beta_1\,\gamma_1, \quad (20)$$

$$x = 0 : \beta_0 + \beta_1\,\gamma_0 + \beta_2. \quad (21)$$

The causal effect definitions clarify this choice. The first alternative is PNIE and the second alternative is TNIE. These two indirect effects were referred to earlier in the continuous outcome case, where it was clear that the TNIE alternative given in Equation 5 represents the total indirect effect in that it includes the interaction effect. The inclusion of the direct effect parameter β_2 in the TNIE comparison of Equations 20 and 21 perhaps surprising at first, but follows clearly from the counterfactual concepts underlying the causal effect definitions.

Continuous *M*, Count *Y*

With a count *Y*, the equation for Y shown in Equation 1 refers to the log rate of the count variable, where the rate is the mean of the count variable. An indirect effect such as $\beta_1 \times \gamma_1$ refers to this log rate as the dependent variable, analogous to y^* in the binary case just discussed. The causal effect definitions instead consider the expected value of the observed Y variable; that is, the rate. Just as with the binary case, two parameters is not sufficient to describe this rate, but all the parameters of the model are involved. This means that the indirect effect should not be computed simply by exponentiating $\beta_1 \times \gamma_1$. The causal effect formulas for the indirect effect with a count variable are shown, for example, in Muthén (2011).

The Poisson and negative binomial models for counts have the same rate (mean) and therefore the same effect formulas. Zero-inflated models need to take into account that the mean is the rate multiplied by $1 - \pi$, where π is the probability of being in the zero class.

With this discussion of intuitively derived effects as a background, the next section presents the causal effect definitions. Following this, examples are discussed.

CAUSAL EFFECT DEFINITIONS

Let $Y_i(x)$ denote the potential outcome that would have been observed for subject i had the treatment variable X been set at the value x, where x is 0 or 1 in this case (this will be generalized to a continuous X variable later on). The $Y_i(x)$ outcome might not be the outcome that is observed for the subject and is therefore possibly counterfactual. For instance, $Y_i(0)$ is counterfactual if subject i received treatment so that $x = 1$ instead of 0. The causal effect of treatment for subject i can be seen as $Y_i(1) - Y_i(0)$, but is clearly not identified given that a subject only experiences one of the two treatments. The average (expected) effect $E[Y(1) - Y(0)]$ is, however, identifiable. Expanding the notation, let $Y(x, m)$ denote the potential outcome that would have been observed if the status of the treatment variable X for subject i was x and the value of the mediator M was m.

Direct and Indirect Effects

The controlled direct effect (CDE) considers a specific fixed value $M = m$ and is defined as

$$CDE(m) = E[Y(1, m) - Y(0, m)|]. \quad (22)$$

The first, index of the first term is 1 corresponding to the treatment group and the first index of the second term is 0 corresponding to the control group. The direct effect notion is motivated by the fact that the effect of treatment on Y is not transferred via M given that its value is fixed. The CDE is perhaps less useful in typical mediation studies in

behavioral sciences but provides a pedagogical starting point as a contrast with other effects. Quoting VanderWeele and Vansteelandt (2009, p. 459):

> While controlled direct effects are often of greater interest in policy evaluation (Pearl, 2001; Robins, 2003), natural direct and indirect effects may be of greater interest in evaluating the action of various mechanisms (Robins, 2003; Joffe et al., 2007).

In contrast to the CDE, the PNDE does not hold the mediator constant, but instead allows the mediator to vary over subjects in the natural way it would vary if the subjects were given the control condition. For example, $Y_i(0, M_i(0))$ is the Y value for subject i when the mediator M for subject i obtains the value it would when $X = 0$. The PNDE is expressed as

$$PNDE = E[Y(1, M(0)) - Y(0, M(0))]. \quad (23)$$

A simple way to view this is to note that Y's first argument, that is x, changes values, but the second does not, implying that Y is influenced by X only directly. The choice of the control condition for M as opposed to the treatment condition is motivated as follows by VanderWeele and Vansteelandt (2009, p. 459) using the term *exposure* to refer to X:

> The pure natural direct effect – expresses the effect that would be realized if the exposure were administered, but its effect on the mediator were somehow blocked, or equivalently, if the mediator were kept at the level it would have taken in the absence of the exposure.

The TNDE considers M at the treatment condition instead of the control condition,

$$TNDE = E[Y(1, M(1)) - Y(0, M(1))]. \quad (24)$$

The TNIE is defined as

$$TNIE = E[Y(1, M(1)) - Y(1, M(0))]. \quad (25)$$

A simple way to view this is to note that the first argument of Y does not change, but the second does, implying that Y is influenced by X due to its influence on M.

The PNIE instead considers $X = 0$,

$$PNIE = E[Y(0, M(1)) - Y(0, M(0))]. \quad (26)$$

The total effect (TE) is

$$TE = E[Y(1) - Y(0)] \quad (27)$$

$$= E[Y(1, M(1)) - Y(0, M(0))]. \quad (28)$$

A simple way to view this is to note that both indices are 1 in the first term and 0 in the second term. In other words, the treatment effect on Y comes both directly and indirectly due to M. The total effect is the sum of direct and indirect effects with two alternative decompositions,

$$TE = PNDE + TNIE = TNDE + PNIE. \quad (29)$$

As mentioned earlier, applications in this article consider the PNDE direct effect, the TNIE indirect effect, and the $TE = PNDE + TNIE$ decomposition.

It is important to note that the counterfactual definitions of effects rely on a set of strong assumptions. These are described in the following quote from Valeri and VanderWeele (2013, p. 140):

> In summary, controlled direct effects require (a) no unmeasured treatment–outcome confounding and (b) no unmeasured mediator–outcome confounding. Natural direct and indirect effects require these assumptions and also (c) no unmeasured treatment–mediator confounding and (d) no mediator–Outcome confounder affected by treatment. It is important to note that randomizing the treatment is not enough to rule out confounding issues in mediation analysis. This is because randomization of the treatment rules out the problem of treatment–outcome and treatment–mediator confounding but does not guarantee that the assumption of no confounding of mediator–outcome relationship holds. This is because even if the treatment is randomized, the mediator generally will not be.

Applying the Causal Effect Definitions

The causal effects are expressed in a general way using expectations and can be applied to many different settings. The X variable need not be binary 0/1, but can be continuous where two different values are compared. Covariates can be included and the effects expressed conditional on certain values of these covariates or summed over them. Moderator variables can be included. Cases studied in the literature include:

- Continuous M, continuous Y (gives the usual SEM formulas unless treatment–mediator interaction).
- Continuous M, categorical Y.
- Categorical M, continuous Y.
- Categorical M, categorical Y.
- Count Y.
- Nominal M, continuous Y.
- Nominal Y.
- Survival Y.

The Appendix shows how the general causal effects expressions are applied to some key types of variables and moderator situations. Further cases are shown in VanderWeele and Vansteelandt (2009) and Muthén (2011).

The next two section's give summaries of two main cases, relating back to the section on intuitive understanding of the effects.

Continuous Y with treatment–mediator interaction. The central component of the general causal effects definitions is the expectation $E[Y(x, M(x^*)]$. Considering this expectation for the model considered earlier with a continuous M and a continuous Y,

$$y_i = \beta_0 + \beta_1\ m_i + \beta_2\ x_i + \beta_3\ x_i\ m_i + \epsilon_{1i}, \tag{30}$$

$$m_i = \gamma_0 + \gamma_i\ x_i + \epsilon_{2i}, \tag{31}$$

the Appendix shows that

$$E[Y(x, M(x^*)] \ = \ \beta_0 + \beta_2\, x + \beta_1 + \beta_3\, x(\gamma_0 + \gamma_1\, x^*). \tag{32}$$

The various causal effects are obtained by taking the difference of such terms for different x and x^* combinations. For example, it is seen that Equation 32 gives the direct and indirect effects for a binary X,

$$PNDE = E[Y(1), M(0)] - E[Y(0), M(0)] = \beta_2 + \beta_3\, \gamma_0, \tag{33}$$

$$TNIE = E[Y(1), M(1)] - E[Y(1), M(0)] = \beta_1\ \gamma_1 + \beta_3\ \gamma_1, \tag{34}$$

effects that were discussed earlier. The Appendix gives the effects for a continuous X.

Binary Y. With a binary Y, the expectation is the probability of $Y = 1$. For a binary treatment variable X, continuous M, and binary Y using probit regression, the Appendix shows that the PNDE, TNIE, and TE effects therefore become the probability differences

$$PNDE = \Phi[probit(1, 0)] - \Phi[probit(0, 0)], \tag{35}$$

$$TNIE = \Phi[probit(1, 1)] - \Phi[probit(1, 0)], \tag{36}$$

$$TE = PNDE + TNIE = \Phi[probit(1, 1)] - \Phi[probit(0, 0)], \tag{37}$$

where Φ is the standard normal distribution function and $probit(x, x^*) = probit(Y(x, M(x^*)))$ for the probit function

$$probit(x, x^*) = [\beta_0 + \beta_2\ x + (\beta_1 + \beta_3\ x)(\gamma_0 + \gamma_1\ x^*)]/\sqrt{v(x)}, \tag{38}$$

where the variance $v(x)$ for x, x^* is the residual variance $v(x) = (\beta_1 + \beta_3 x)^2\sigma_2^2 + 1$ shown in Equation 14 and where the parameter names of the analogous continuous Y model of Equation 30 and 31 are used. For example, it is seen how TNIE using Equation 38 with no treatment–mediation interaction ($\beta_3 = 0$) leads to comparing Equation 20 and 21 for a binary X, as was discussed previously.

With a logistic regression instead of probit regression for the binary Y, the effect formulas are not explicit as for probit but the effects are obtained via numerical integration.

LATENT VARIABLES

To the list of causal effect cases shown earlier, this article adds the case of latent X, M, and Y. The general causal effect expressions are applicable also here and the formulas are straightforward because the latent variables are continuous and are related by linear regression. For example, the continuous exposure variable X might be only indirectly measured by a set of indicators, where a latent exposure variable can avoid measurement error in an observed score. As is well known, ignoring measurement error in a predictor leads to biased regression coefficients. Likewise, it might be better to represent the mediator as a latent variable measured by multiple indicators. The consequences of measurement error in the mediator can be severe as has been shown in Hoyle and Kenny (1999) using a traditional mediation model with continuous dependent variables; that is, considering regular SEM effects. Hoyle and Kenny (1999) made a comparison using a fallible mediator versus using a latent variable mediator measured by multiple fallible indicators. The trade-off between accuracy and precision when using latent variable mediation models was studied in Ledgerwood and Shrout (2011). The topic of measurement error in the mediator has also been discussed in the casual inference literature by le Cressie, Debeij, Rosendaal, Cannegieter, and Vandenbroucke (2012) and VanderWeele, Valeri, and Ogbum (2012) considering causal effects for a binary outcome. In all these papers, it is shown that measurement error in the mediator causes an underestimated indirect effect and an overestimated direct effect in line with regression on one predictor measured with error and the other without error (Carroll, Rupport, Stefanski, & Crainiccanu, 2006). This article (contributes to the topic by focusing on causal effects for a binary outcome, comparing the use of a fallible mediator with the use of a latent variable measured by multiple fallible indicators. The SEM literature has not considered causal effects for a binary Y and the latent variable solution has not been studied in the causal inference literature.

The use of a latent variable mediator might also serve to ameliorate the normality assumption of the mediator residual. For instance, using as a mediator the sum of a set of skewed binary indicators results in a nonnormal distribution whereas a nonnormal factor is not necessitated by having such binary indicators.

Monte Carlo Simulations for a Mediator Measured With Error

A Monte Carlo study is carried out to study the effects of measurement error in a continuous mediator on the causal indirect and direct effects for a binary outcome. This means that the effects are expressed in a probability metric comparing two values of the exposure, which is in this case a binary variable. A comparison is made between using as the mediator a single fallible indicator, a sum of similarly fallible indicators, and a latent variable measured by these fallible indicators using a standard factor analysis model. Figure 4 shows the latent variable situation. Although biases can in principle be derived analytically for population values, it is of interest to study the sampling behavior for a limited sample size. A sample size of $n = 200$ is chosen as representative of many behavioral science intervention studies.

The mediator measurement error is studied in terms of reliability levels of 0.6 and 0.8 (the reliability of blood pressure readings has been estimated to be 0.67; see Shepard, 1981; and a reliability of 0.7–0.8 is often seen for good items in behavioral science measurement instrument). The parameter values are shown in the M*plus* input given in the Appendix. The exposure has a positive effect on the mediator and a negative effect on the outcome. The mediator has a negative effect on the outcome. The TNIE is –0.161 and the PNDE is –0.132. These numbers represent the drop in probability of $Y = 1$ for subjects randomized to treatment versus control group. This means that the total effect is –0.293; that is, the difference in probability of $Y = 1$ comparing treatment group to control group. Because the probability is on a 0 to 1 metric, this might be regarded as a moderate effect size. The number of mediator indicators is varied as 1, 3, and 6. Maximum-likelihood estimation with a sample size of $n = 200$ is used together with 500 replications.

Table 1 shows the results of the Monte Carlo simulations. The coverage is excellent for all parameters and effects using the latent variable modeling with is multiple indicators and is not reported. Table 1 gives the average and standard deviation of the causal effects.

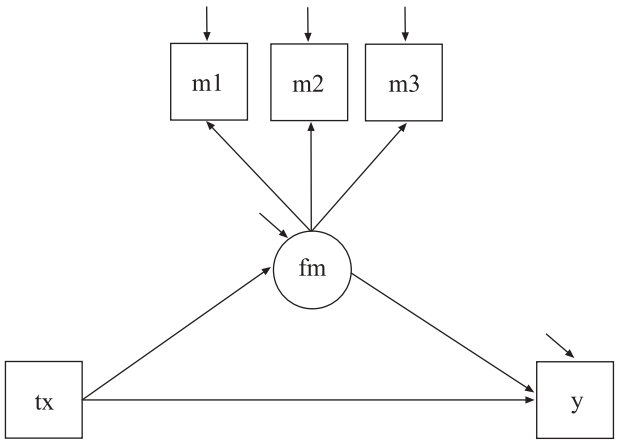

FIGURE 4 A mediation model for a Monte Carlo study of measurement error in the mediator with binary outcome.

Consider first the results for using a single item or sums of items as the mediator. It is seen that using a single item with reliability 0.6 as the mediator causes large biases. For example, the indirect effect (TNIE) is only about half the size of the true value, –0.085 versus –0.161. Furthermore, the relative size of the indirect effect (TNIE) and the direct effect (PNDE) is reversed so that the indirect effect is now less than half of the direct effect, –0.085 versus –0.210, as compared to the true values of –0.161 and –0.132. Summing three such items improves on the biases only a little and summing six such items is still not quite sufficient at this reliability level. For item reliability of 0.8 the sum of six items approaches the true values for the effects and correctly orders the size of the indirect and direct effects. The biases seen in Table 1 are largely unchanged by using larger sample sizes than the $n = 200$ studied here.

The latent variable model gives very small biases including the case with reliability 0.6 and using only three indicators of the latent variable. A single-indicator case would call for a known reliability and is not studied here. The standard deviation of the effect estimates for the latent variable approach is slightly larger than for the approach of using a sum, but this is a negligible trade-off for the much smaller bias (Ledgerwood & Shrout, 2011, reported larger trade-offs in their settings).

TABLE 1
Monte Carlo Study of Measurement Error in the Mediator With Binary Outcome: Average (Standard Deviation) of Indirect (TNIE) and Direct (PNDE) Effects in Probability Metric Over 500 Replications With $n = 200$

		Number of Items		
Effect	*Approach*	*1*	*3*	*6*
Item reliability = 0.6				
TNIE (true value = –0.161)	Sum	−0.085	−0.125	−0.140
		(0.030)	(0.035)	(0.037)
	Latent		−0.163	−0.161
			(0.047)	(0.043)
PNDE (true value = –0.132)	Sum	−0.210	−0.169	−0.152
		(0.059)	(0.055)	(0.049)
	Latent		−0.131	−0.131
			(0.058)	(0.050)
Item reliability = 0.8				
TNIE (true value = –0.161)	Sum	−0.121	−0.145	−0.153
		(0.034)	(0.037)	(0.039)
	Latent		−0.161	−0.161
			(0.041)	(0.042)
PNDE (true value = –0.132)	Sum	−0.176	−0.151	−0.140
		(0.054)	(0.051)	(0.048)
	Latent		−0.135	−0.131
			(0.052)	(0.049)

Note. The sum approach uses as the mediator the sum of the items and the latent approach uses as the mediator a latent variable measured by the items. TNIE = total natural indirect effect; PNDE = pure natural direct effect.

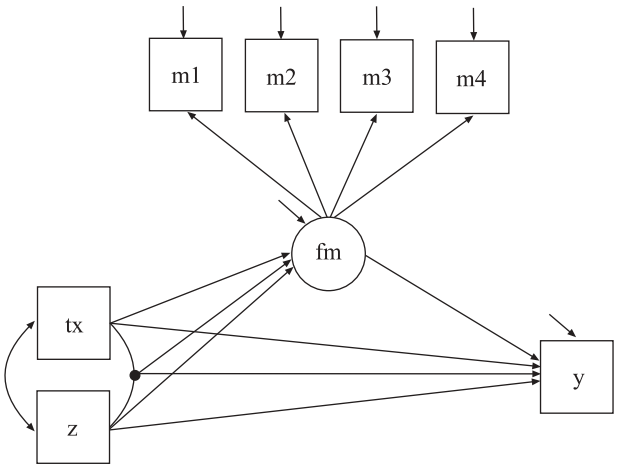

FIGURE 5 A moderated mediation model for an intervention study of aggressive behavior in the classroom and juvenile court record.

An Example: An Intervention Study of Aggressive Behavior in the Classroom and Juvenile Court Record

Data for this example are from a randomized field experiment in Baltimore public schools where a classroom-based intervention was aimed at reducing aggressive, disruptive behavior among elementary school students (Ialongo et al., 1999). Figure 5 shows the mediation model. The ultimate outcome is a binary variable indicating whether or not the student obtained a juvenile court record by age 18 or an adult criminal record. The mediator is a continuous latent variable, the factor labelled *fm* in Figure 5, which is measured in Grade 5 by the four items Harms Others, Breaks Things, Takes Property, and Fights, labeled $m1$ through – $m4$ in Figure 5. These items are rated by teachers on a scale from 1 (*almost never*) through 6 (*almost always*). The intervention started in the fall of Grade 1. This methods illustration ignores the fact that the intervention is classroom based. A treatment–baseline interaction is hypothesized so that initially more aggressive children might benefit more from the intervention. Corresponding to this, a moderator variable is added in the form of an aggressive behavior score recorded in the fall of Grade 1 before the intervention started. The interaction between this variable and the binary intervention variable is assumed to influence both the mediator and the ultimate outcome. The analysis to be presented involves $n = 250$ boys in treatment and control classrooms with complete data. A juvenile court record by age 18 or an adult criminal record is observed for 50% of the sample. It could be noted that in these analyses a significant treatment–mediator interaction effect is not found and is excluded from the model.

Two analyses are presented, both using maximum-likelihood estimation. In the first analysis, the mediator is an observed variable formed as the sum of the four aggressive behavior items. In the second analysis the factor behind the items is the mediator, forming a measurement model that takes into account potential measurement error of the items. The difference in mediation results illustrates the consequence of ignoring measurement error. The analysis focuses on the PNDE and TNIE defined earlier, presented for different values of the moderator variable varying from one standard deviation below its mean to one standard deviation above its mean. The M*plus* input for both analyses are shown in the Appendix.

The direct effect PNDE is not significant in either of the two analyses, whereas the indirect effect TNIE is significant in both. The choice of probit versus logit link makes almost no difference in the effect estimates. The use of regular maximum likelihood with symmetric confidence intervals or bootstrap standard errors and nonsymmetric confidence intervals also makes almost no difference. For the analysis using the mediator formed as the sum of the four aggressive behavior items, TNIE = –0.046 with z test value –2.101 when computed at the mean of the moderator. This implies a significant indirect effect reduction by 0.046 in the probability of a juvenile court record by age 18 or an adult criminal record. The size of the effect can be related to the approximately .5 probability observed in the sample. The corresponding result for the model using the latent mediator is a somewhat larger indirect effect, TNIE = –0.053 with z test value –2.144. A plot of the TNIE effect for the latent mediator model is seen in Figure 6. The plot indicates that the indirect effect has an increasingly large negative value when the moderator value increases and is significant for a range of moderator values from –0.45 to 0.65 (the mean of the moderator is zero and its standard deviation is approximately one). The results are commensurate with the hypothesis that initially more aggressive children benefit more from the intervention.

It could be noted that standardization of the effect estimates is not needed. The effects are on a probability scale and are therefore in a sense already standardized with respect to this dependent variable. Also, the treatment variable is binary so there is no need for standardization with respect to this variable either. For a continuous X variable there is also no need for standardization with respect to X. The two X values for which the effect is considered can take into account the mean and standard deviation of X and choose suitable values accordingly.

DISCUSSION

This article shows that the causal inference literature provides important lessons for structural equation modelers. The counterfactual reasoning that the causal effects are built on provides a clear understanding of how the effects should be defined and provides a general approach for deriving them. With the effects being easily accessible in various software packages, it is time for these methods to be explored in SEM.

This article also shows how important it is to take measurement error into account in estimating the effects. In particular, measurement error in the mediator when ignored can give a strong distortion of indirect and direct

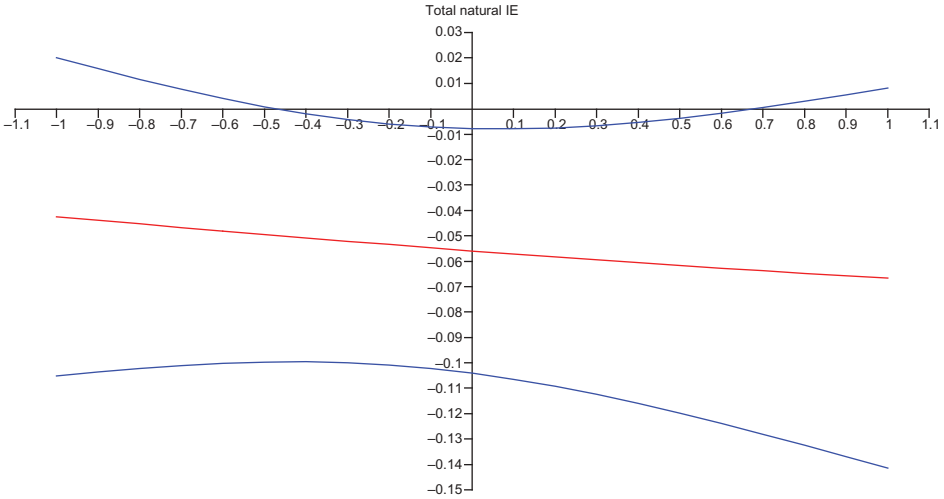

FIGURE 6 Total natural indirect effect (TNIE) as moderated by baseline aggressive behavior score for an intervention study of aggressive behavior in the classroom and juvenile court record.

effects. Using a factor analysis model for multiple indicators of the latent variable mediator is a useful way to avoid such distortions.

Several other mediation cases have been omitted from this article for the sake of brevity. One case is that of a binary mediator, which could be combined with a continuous, binary, or count outcome. With a binary mediator the integration over the mediator used in the derivation of the general mediation formulas is replaced by a sum over the two categories (see, e.g., VanderWeele & Vansteelandt, 2009). A second case is that of a nominal mediator discussed in Muthén (2011). This can also be generalized to a latent class mediator. A third case is that of a polytomous, ordered (ordinal) outcome, where the expectation in the general mediation formulas could refer to a particular outcome category or sets of categories. As the number of categories of an ordinal variable increases one might also consider the effects for an underlying latent response variable to be a more relevant effect summary. Muthén (2011) discussed the use of causal effects for a latent response variable M^* in the case of an ordinal mediator.

The causal inference literature stresses that the derivations of the causal effects rely on strong assumptions. The study of sensitivity to the underlying assumptions has been proposed by Imai, Keele, and Tingley (2010) and VanderWeele (2010). Muthén (2011) showed how the results of the Imai et al. sensitivity analysis can be visualized in M*plus*.

REFERENCES

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

Carrol, R. J., Rupper, D., Stefanski, L. A., & Crainiceanu, C. (2006). *Measurement error in nonlinear models*. (2nd ed). Boca Raton, FL: Chapman & Hall.

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.

Ialongo, L. N., Werthamer, S., Kellam, S. K., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression and antisocial behavior. *American Journal of Community Psychology*, *27*, 599–641.

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

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, *69*, 307–321.

Joffe, M., Small, D., & Hsu, C.-Y. (2007). Defining and estimating intervention effects for groups that will develop an auxiliary outcome. *Statistical Science*, *22*, 74–97.

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

Kraemer, H. K., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology*, *27*, S101–S108.

le Cressie, S., Debeij, J., Rosendaal, F. R., Cannegieter, S. C., & Vandenbroucke, J. P. (2012). Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. *Epidemiology*, *23*, 551–560.

Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology*, *101*, 1–15.

MacKinnon, D. P. (2008). *An introduction to statistical mediation analysis*. New York, NY: Erlbaum.

MacKinnon, D. P. & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, *17*, 144–158.

MacKinnon, D. P., Lockwood, C. M., Brown, C. H., Wang, W., & Hoffman, J. M. (2007). The intermediate endpoint effect in logistic and probit regression. *Clinical Trials*, *4*, 499–513.

Muthén, B. (1979). A structural probit model with latent variables. *Journal of the American Statistical Association*, *74*, 807–811.

Muthén, B. (2011). *Applications of causally defined direct and indirect effects in mediation analysis using SEM in M*plus (Technical Report). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998–2012). *M*plus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

Pearl, J. (2001). Direct and indirect effects. In J. Breese & D. Koller (Eds.), *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*. pp. 411–420. San Francisco, CA: Morgan Kaufman.

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

Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In P. Greeen, N. L. Hjort, & S. Richardson (Eds.), In *Highly structured stochastic systems*. (pp. 70–81). New York, NY: Oxford University Press.

Shepard, D. S. (1981). Reliability of blood pressure measurements: Implications for designing and evaluating programs to control hypertension. *Journal of Chronic Diseases*, *34*, 191–209.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2013). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*, 1–38.

Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*, 137–150.

VanderWeele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*, *21*, 540–551.

VanderWeele, T. J., Valeri, L., & Ogburn, E. L. (2012). The role of measurement error and misclassification in mediation analysis. *Epidemiology*, *23*. 561–564.

VanderWeele, T. J., & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, *2*, 457–468.

APPENDIX

Causal Effect Formulas

The causal effect formulas found in the literature (see, e.g., Imai et al., 2010; Pearl, 2001; VanderWeele & Vansteelandt, 2009) are summarized here for the article to be self-contained. This also has the advantage that the formulas are expressed in the notation used in this article and in the M*plus* program. Derivations are given of the casual effects for both continuous Y and binary Y for the following case: a continuous mediator, a treatment–mediator interaction, an exogenous moderator Z, and a covariate C. The effects are as follows for the exposure $X = x$ compared to $X = x^*$ (for a binary X, $x = 1$ and $x^* = 0$, for say treatment versus control group):

$$TNIE = E[Y(x, M(x))|C = c, Z = z] - E[Y(x, M(x^*))|C = c, Z = z], \tag{A.1}$$

$$PNDE = E[Y(x, M(x^*))|C = c, Z = z] - E[Y(x^*, M(x^*))|C = c, Z = z], \tag{A.2}$$

$$PNIE = E[Y(x^*, M(x))|C = c, Z = z] - E[Y(x^*, M(x^*))|C = c, Z = z], \tag{A.3}$$

$$TNDE = E[Y(x, M(x))|C = c, Z = z] - E[Y(x^*, M(x))|C = c, Z = z], \tag{A.4}$$

$$TE = E[Y(x, M(x))|C = c, Z = z] - E[Y(x^*, M(x^*))|C = c, Z = z], \tag{A.5}$$

Imai et al. (2010) referred to the indirect effects as average causal mediated effects with TNIE and PNIE translated as ACME (treated) and ACME (control). Imai et al. referred to the direct effects as average direct effects with TNDE and PNDE translated as ADE (treatment) and ADE (control).

Using the fact that for a joint distribution of Y, V the expectation in the marginal distribution of Y is

$$E(Y) = \int_{-\infty}^{+\infty} E(Y|V) \times f(V)\, \partial V, \tag{A.6}$$

the key component of the causal effect definitions, $E[Y(x, M(x^*))|C = c, Z = z]$, can be expressed as follows integrating over the mediator M,

$$E[Y(x, M(x^*))|C = c, Z = z] = \int_{-\infty}^{+\infty} E[Y|C = c, Z = z, X = x, M = m] \times f(M; E[M|C = c, Z = z, X = x^*)\, \partial M. \tag{A.7}$$

Continuous outcome. Consider first the model for a continuous outcome,

$$y_i = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, m_i + \beta_4\, \epsilon_i + \beta_5\, z_i + \beta_6\, x_i\, z_i + \epsilon_{1i}, \tag{A.8}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \gamma_3\, z_i + \gamma_4\, x_i\, z_i + \varepsilon_{2i}. \tag{A.9}$$

Using Equation A.7, the model gives

$$E[Y(x, M(x^*))|C = c, Z = z] = \tag{A.10}$$

$$= \beta_0 + \beta_2\, x + \beta_4\, c + \beta_5\, z + \beta_6\, x\, z + (\beta_1 + \beta_3\, x) \int_{-\infty}^{+\infty} f(m; \gamma_0 + \gamma_1\, x^* + \gamma_2\, c + \gamma_3\, z + \gamma_4\, x^* z, \sigma_2^2)\, \partial M, \tag{A.11}$$

$$= \beta_0 + \beta_2\, x + \beta_4\, c + \beta_5\, z + \beta_6\, x\, z \tag{A.12}$$

$$+ (\beta_1 + \beta_3\, x)(\gamma_0 + \gamma_1\, x^* + \gamma_2\, c + \gamma_3\, z + \gamma_4\, x^*\, z). \tag{A.13}$$

Using different combinations of x, x^* values in Equations A.12 and A.13, the different causal effects are obtained using Equations A.1 through A.5. For example,

$$TNIE = (\beta_1 + \beta_3\, x)(\gamma_1 + \gamma_4\, z)(x - x^*), \tag{A.14}$$

$$PNDE = (\beta_2 + \beta_6\, z + \beta_3\, (\gamma_0 + \gamma_1\, x^* + \gamma_2\, c + \gamma_3\, z + \gamma_4 x^* z))(x - x^*). \tag{A.15}$$

Binary outcomes. Consider next the model for a binary outcome so that Equation A.8 refers to a logit or probit regression. Noting that $E(Y) = P(Y = 1)$, the formulas for probit are obtained by once again integrating over M,

$$E[Y(x, M(x^*))|C = c, Z = z] = \tag{A.16}$$

$$= \int_{-\infty}^{\infty} E[Y|C=c, Z=z, X=x, M=m] \times f(M|C=c, X=x^*)\partial M \tag{A.17}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{probit(y)} f(z;0,1)\partial z \times f(M; \gamma_0 + \gamma_1\ x^* + \gamma_2\ c + \gamma_3\ z + \gamma_4\ x^* z,\ \sigma_2^2)\ \partial M \tag{A.18}$$

$$= \int_{-\infty}^{probit(x,x*)} f(z;0,1)\partial z = \Phi(probit(x, x^*)), \tag{A.19}$$

where Φ is the standard normal distribution function and the last equality can be derived by a variable transformation and a change of the order of integration as in Muthén (1979; p. 810, Appendix Theorem) with

$$probit(x, x^*) = [\beta_0 + \beta_2\ x + \beta_4\ c + \beta_5\ z + \beta_6\ xz + (\beta_1 + \beta_3 x)\ (\gamma_0 + \gamma_1\ x^* + \gamma_2\ c + \gamma_3\ z + \gamma_4\ x^* z)]/\sqrt{v(x)}, \tag{A.20}$$

where the variance $v(x) = (\beta_1 + \beta_3 x)^2 \sigma_2^2 + 1$, where σ_2^2 is the residual variance for the continuous mediator. With logit, the integration over M in Equation A.18 needs to be carried out using numerical integration.[1] Applying Equation A.1 through A.5 with Equations A.19 and A.20 gives the causal effects for probit with the special cases shown earlier for a binary X.

M*plus* Scripts

Muthén (2011) showed how causal effects can be computed using the M*plus* command MODEL CONSTRAINT. This approach, however, requires that the user specifies the formulas for the effects, which is cumbersome. M*plus* Version 7.2 does this work behind the scenes. Following is the M*plus* Version 7.2 language and the different mediation and moderation cases considered for causal effects defined via counterfactuals, including which variables can be specified as latent. The modeling is currently restricted to a single mediator.

1. No moderation:
   - Y IND M X.
   - All 3 can be latent.
2. Moderation with X*M (3 arguments after MOD):
   - y MOD M XM X;
   - Y can be latent
3. Moderation with Z
   - Involving X and M (5 arguments after MOD):
     - Y MOD M Z(low, high, increment) MZ XZ X;
     - Only Y can be latent
   - Involving M and not X (4 arguments after MOD):
     - Y MOD M Z(low, high, increment) MZ X;
     - X and Y can be latent
   - Involving X and not M (4 arguments after MOD):
     - Y MOD M Z(low, high, increment) XZ X;
     - M and Y can be latent

For a continuous X, the two X values (exposure levels) that, the effects compare are given in parentheses after X: $X(x\ x^*)$. If these two values are not mentioned, a binary X is assumed corresponding to comparing treatment ($X = 1$) to control ($X = 0$): $X(1\ 0)$. For controlled direct effects an M value is placed in parenthesis: $M(m)$. For moderation with Z, a "Loop" plot is generated via the PLOT command, showing the estimated effects and their confidence intervals.

[1] In the logit case, M*plus* carries out numerical integration with 101 integration points.

TABLE A.1
M*plus* Input for the Step 1 Internal Monte Carlo Analysis Using a Latent Variable Model With Three Indicators

TITLE:	Simulating binary X, cont latent M, binary Y Step 1
MONTECARLO:	NAMES = y m1-m3 x; GENERATE = y(1 p); CATEGORICAL = y; NOBSERVATIONS = 200; NREPS = 500; REPSAVE = all; SAVE = n200Perc20rep*.dat; CUTPOINTS = x(0);
MODEL POPULATION:	x@1; fm BY m1-m3*1; fm*1; m1-m3*.67; !reliability 0.6 y ON x*-1 fm*-2.5; [y$1*.75]; fm ON x*.7; !R-square 0.10 (binary x)
ANALYSIS:	ESTIMATOR = ML; LINK = PROBIT;
MODEL:	fm BY m1-m3*1; fm@1; m1-m3*.67; !reliability 0.6 y ON x*-1 fm*-2.5; [y$1*.75]; fm ON x*.7; !R-square 0.10 (binary x)
MODEL INDIRECT:	y IND fm x;

M*plus* Input for the Monte Carlo study.

Table A.1 gives the M*plus* input for the Step 1 internal Monte Carlo analysis. This step uses the latent variable model with three indicators. This step also saves the generated data for a Step 2 external Monte Carlo analysis where the indicators are summed. For a general description of Monte Carlo simulations using M*plus*, see Chapter 12 of the Version 7 User's Guide (Muthén & Muthén, 1998–2012).

Table A.2 gives the M*plus* input for the Step 2 external Monte Carlo analysis. This step uses the sum of the three indicators as the mediator.

TABLE A.2
M*plus* Input for the Step 2 External Monte Carlo Analysis Using a Sum of Indicators as the Mediator

TITLE:	Simulating binary X, cont latent M, binary Y Step 2
DATA:	FILE = n200Perc20replist.dat; TYPE = MONTECARLO;
VARIABLE:	NAMES = y ml-m3 x; USEVARIABLES = y x SUM; CATEGORICAL = y;
DEFINE:	SUM = SUM(ml-m3);
ANALYSIS:	ESTIMATOR = ML; LINK = PROBIT;
MODEL:	y ON x*-.2 SUM*-.25; [y$l*.75]; SUM ON x*.7; !R-square 0.10
MODEL INDIRECT:	y IND SUM x;

M*plus* Input for aggressive behavior analysis.

Table A.3 shows the input for the model where the mediator is formed as the sum of the four aggressive behavior items. The *Z* variable is centered and has a standard deviation of approximately 1. The moderation effect is studied for *Z* values ranging from –1 to +1, which is approximately 1 *SD* below to 1 *SD* above the *Z* mean.

TABLE A.3
M*plus* Input for Aggressive Behavior Mediation Model With a Moderator *Z* and an Observed Mediator *M*

```
VARIABLE:
            . . .
            USEVARIABLES = y tx xz m z;
            CATEGORICAL = y;
            USEOBSERVATIONS = gender EQ 1 AND
              (desgn11s EQ 1 OR desgn11s
            EQ 2 OR desgn11s EQ 3 OR desgn11s EQ 4);
DEFINE:     IF(desgn11s EQ 4)THEN tx = l;
            IF(desgn11s EQ 1 OR desgn11s EQ 2 OR
              desgn11s EQ 3) THEN tx=0;
            y = juvadl;
            m = sum(toc05110 toc05114 toc05121 toc
              05124);
            z = sctaa11f;
            CENTER z(GRANDMEAN);
            xz = tx*z;
ANALYSIS:   ESTIMATOR = MLR;
            LINK = PROBIT;
MODEL:      y ON m z xz tx;
            m ON z xz tx;
MODEL INDIRECT: y MOD m z(-l, 1,0.1) xz tx;
OUTPUT:     TECH1 TECH8 SAMPSTAT;
PLOT:       TYPE = PLOT3;
```

Table A.4 shows the input for the model where the mediator is formed as a factor measured by four aggressive behavior items.

TABLE A.4
M*plus* Input for Aggressive Behavior Mediation Model With a Moderator *Z* and a Latent Mediator fm Measured by Four Items

```
VARIABLE:
            . . .
            USEVARIABLES = toc05110 toc05114 toc05121
              toc05124 y tx z xz;
            CATEGORICAL = y;
            USEOBSERVATIONS = gender EQ 1 AND
              (desgn11s EQ 1 OR desgn11s
            EQ 2 OR desgn11s EQ 3 OR desgn11s EQ 4);
DEFINE:     IF(desgn11s EQ 4)THEN tx=l;
            IF(desgn11s EQ 1 OR desgn11s EQ 2 OR
              desgn11s EQ 3)THEN tx=0;
            y = juvadl;
            z = sctaa11f;
            CENTER z(GRANDMEAN);
            xz = tx*z;
ANALYSIS:   ESTIMATOR = MLR;
            LINK = PROBIT;
MODEL:      fm BY toc05110-toc05124;
            y ON fm z xz tx;
            fm ON z xz tx;
MODEL INDIRECT: y MOD fm z(–l, 1,0.1) xz tx;
OUTPUT:     TECH1 TECH8 SAMPSTAT;
PLOT:       TYPE = PLOT3;
```