Basic and Applied Social Psychology

ISSN: 0197-3533 (Print) 1532-4834 (Online) Journal homepage: www.tandfonline.com/journals/hbas20

Reversing Arrows in Mediation Models Does Not Distinguish Plausible Models

Felix Thoemmes

To cite this article: Felix Thoemmes (2015) Reversing Arrows in Mediation Models Does Not Distinguish Plausible Models, Basic and Applied Social Psychology, 37:4, 226-234, DOI: 10.1080/01973533.2015.1049351

To link to this article: https://doi.org/10.1080/01973533.2015.1049351

Published online: 04 Aug 2015.

BASIC AND APPLIED SOCIAL PSYCHOLOGY, 37:226–234, 2015

ISSN: 0197-3533 print/1532-4834 online
DOI: 10.1080/01973533.2015.1049351

Reversing Arrows in Mediation Models Does Not Distinguish Plausible Models

Felix Thoemmes
Cornell University

Reversing arrows in the classic tri-variate X-M-Y mediation models as a test to check whether one mediation model is superior to another is inadmissible. Presenting evidence that one tri-variate mediation model yields a significant indirect effect, whereas one with some reversed arrows does not, is not proof or even evidence that one model should be preferred. In fact, the significance of the indirect or any other effect can never be used to infer whether one model should be preferred over another, if the models are in the same so-called equivalence class. The practice of running several mediation models with reversed arrows to decide which model to prefer should be abandoned. The only way to choose among equivalent models is through assumptions that are either fulfilled by design features or invoked based on theory. Similar arguments about reversing arrows in mediation models have been made before, but this current work is the first to derive this result analytically for the complete (Markovian) equivalence class of the tri-variate mediation model.

A mediator M in the tri-variate X-M-Y mediation model cannot be statistically distinguished from a confounder (Fiedler, Schott, & Meiser, 2011; MacKinnon, Krull, & Lockwood, 2000). In fact, as we show, it cannot be distinguished from a common cause, common effect, or a mediator in the reverse direction using statistics alone. However, there is still a persistent belief that it can be helpful to test alternative mediation models by switching arrows and comparing size and significance of the indirect effects. Without wanting to single out specific publications or authors, this practice can be seen in published work; for example, Fredrickson, Tugade, Waugh, and Larkin (2003) used this technique to argue that one mediation model (the one with the significant effect) was more plausible than two alternative models (the ones without significant effects). Anecdotally, the author can also attest that it is not uncommon that reviewers (or editors) request that in addition to a presented mediation model, some arrows should be reversed, and the model tested again, presumably to gain insight into which model is preferred.

The (erroneous) thought behind this practice is that if the mediated effect of one model is significant, but the other model has no significant mediated effect, then the mediation model with the significant indirect effect is more plausible. However, none of the six models that one could form by reversing the three *directed* arrows in the tri-variate mediation model can be preferred on the basis of the significance of any of the resulting effects.[1] The reason behind this inability to prefer one model over the other is that they are all in the same *equivalence class*. The equivalence class is defined as a class of models that all have the same implied covariance matrix. All models in this class have identical fit indicies, and the same support from the data (MacCallum, Wegener, Uchino, & Fabrigar, 1993; Meek, 1995; Stelzl, 1986). One way to check whether models are in the same equivalence class is to enumerate the so-called d-separation constraints (Pearl, 2000). D-separation is a graphical criterion for conditional independence. A brief introduction to d-separation is given in Appendix A. For an additional introduction to d-separation for social

Correspondence should be sent to Felix Thoemmes, Cornell University, MVR G62A, Ithaca, NY 14853. E-mail: felix.thoemmes@cornell.edu

[1]If one allows replacement of directed with bidirected arrows, 27 models can be formed, and again these are statistically indistinguishable from each other, and none can be preferred over the other based on any of the resulting effects.

scientists, see, for example, Hayduk et al. (2003). Models without latent variables or bidirected arrows (also called "Markovian" models) are in the same equivalence class, if and only if the two models have the same d-separation constraints (Pearl, 2000). The classic tri-variate mediation model has in fact no d-separation constraints (see also Appendix A for this result), and therefore any other tri-variate model that has *any* paths (no matter which direction) between all three variables is in the same equivalence class. Models in the same equivalence class, however, may have vastly different path coefficients, and it is this fact that leads researcher to erroneously believe that one model may be superior to another. The following trivial example shows that this is incorrect.

In this example, we consider only fully standardized variables (all variables have mean 0 and variance 1), but our results also hold true if unstandardized variables are used.[2] The complete standardization of the variables and hence coefficients is helpful to easily judge their respective sizes. The true model is shown in Figure 1a. There is a strong direct effect of X on Y ($\gamma' = .70$), a weak effect of X on the mediator M ($\alpha = .10$), and a medium effect of the mediator M on the outcome Y ($\beta = .40$). Using the standard approach (MacKinnon, 2008) of forming a product term of path coefficients to quantify the indirect effect yields $\alpha\beta = .04$ for the true model. Given that this is a completely standardized coefficient, its size is rather small and it probably will not be significant, except in very large samples. Assuming this true model, we can derive standardized path coefficients of an incorrect model, in which the roles of M and Y have been exchanged through reversal of the arrow between M and Y. Derivations are presented in Appendix B. In this model, the direct effect of X on Y is negative ($\gamma'^{\star} = -.55$), the effect of X on M is large ($\alpha^{\star} = .74$), and the effect of M on Y even larger ($\beta^{\star} = .88$). Consequently the indirect effect is large, $\alpha^{\star}\beta^{\star} = .65$, and would presumably become significant in all but the smallest samples. A researcher who uses the magnitude and/or significance of the indirect effect (or any other path coefficient) to determine which model is more supported by the data would be misled. In this example, it is the true model that has the smaller, presumably nonsignificant indirect effect and the incorrect model that shows a large indirect effect. This example is one of an infinite number of situations in which this reversal of significant and insignificant results between the true and other incorrect models could occur. To show that this simple example is not a cherry-picked arrangement of path coefficients but that this is a

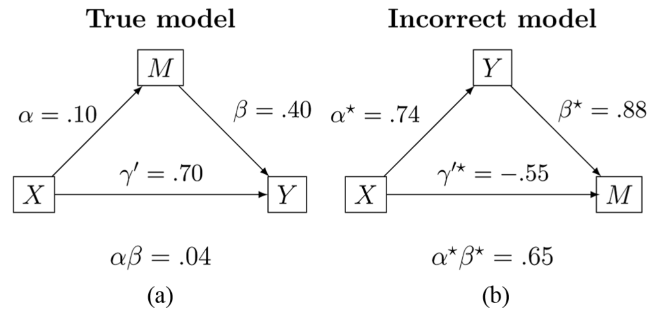

FIGURE 1 True and incorrect model, showing standardized path coefficients.

persistent problem no matter the true underlying mediation model, we conducted an analytic study that explored the asymptotic behavior of path coefficients in a very large number of models.

ANALYTIC STUDY

In our analytic study, we wanted to explore how the magnitude of the indirect effect changes across alternative models. Specifically, we were interested in which alternative mediation model would yield the largest indirect effect, assuming a particular true model. Given that applied researchers use the argument that the model with the larger indirect effect is the correct one, we wanted to see how often a strategy of comparing indirect effects across models and then picking the one with the large indirect effect actually yields the correct model. Given an assumed true model, we derive all asymptotic path coefficients and asymptotic indirect effects of alternative models and record these path coefficients. Details on the mathematical derivation of these effects are found in Appendix B. We then compare whether the model with the largest effect is in fact the true mediation model. This process mimics that of a researcher who has data on variables X, M, Y, and probes different mediation models to make an argument that the one with the large indirect effect is the true one.

Assuming a true model with three variables, X, M, Y, and path coefficients α, β, and γ', as depicted in Figure 2a, we can rearrange arrows in a total of five additional configurations (b to f), assuming only direct effects between variables, and excluding a model with cycles.

Each of these models makes vastly different claims about directions of effects. In some of these models, M is a mediator in the opposite direction (b) of the true model, a common effect (c and d), or a common cause (e and f). Differently said, in some models (d and e) X is the mediator of a relationship between M and Y, and in other models (c and f) Y is the mediator of a

[2] The advantage of using standardized variables is that some of the underlying algebra simplifies and that path coefficients are automatically completely standardized.

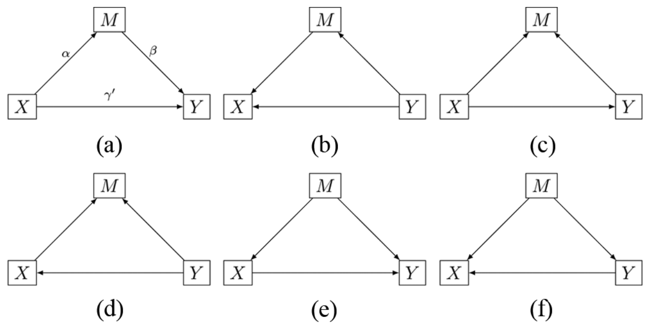

FIGURE 2 Six alternative mediation models that are all in the same equivalence class but may have vastly different path coefficients.

relationship between X and M. Of importance, all six models are in the same equivalence class and cannot be distinguished based on statistical evidence, despite the fact that path coefficients are not identical across models.

Given the true model in Figure 2a, we can derive an asymptotic correlation matrix based on the given path coefficients, and from this matrix we can derive all implied regression coefficients (see Appendix B for details on derivations). There is no need to simulate any data, as all necessary statistics can be analytically derived from the ground truth. To explore a vast parameter space of path coefficients, we first set up vectors for every path coefficient α, β, and γ' from model (a) in Figure 2 that ranged from -1.5 to 1.5 (note that standardized path coefficients, unlike correlations, can go beyond 1) in increments of .05, for a total of 61 unique values for each path coefficient. Then we formed every possible combination of these path coefficients, for a total of $61^3 = 226{,}981$ unique models. We then derived asymptotic correlation matrices for each of these combinations. In a next step, we checked whether these correlation matrices were positive definite. It was expected that many models would yield nonpositive definite matrices as we varied the path coefficients without any constraints over a very large parameter space. Nonpositive definite matrices occur every time when the path coefficients imply correlations that go above 1 or below -1. After nonpositive definite matrices were excluded, 60,447 combinations remained. To visualize the parameter space that we covered, Figure 3 presents histograms of all correlations, and path coefficients. Scatterplots of these estimates are not shown but were inspected and show that the full range of combinations of correlations and path coefficients was captured. In summary, the parameter space that was covered in the study was practically exhaustive of *every* possible correlational pattern of three variables that can be generated from the true model in Figure 2a.

In a next step, we derived asymptotic regression coefficients for every of the six possible mediation models, based on the implied correlation matrix. Then we computed the indirect effect as the product of regression coefficients and simply evaluated which model showed the largest absolute indirect effect. To replicate our study, we provide complete R code for the analysis in Appendix C. We also provide R code online at `www.human.cornell.edu/hd/qml`. The R code is set up in a way that running it replicates all of our results, except graphs.

RESULTS

Evaluating the asymptotic indirect effect of all models over the complete parameter space revealed that indirect effects varied greatly. Every incorrect model yielded indirect effects that were comparable in magnitude with the true indirect effect. Histograms of all indirect effects are shown in Figure 4. What is even more important, however, is the fact that the true indirect effect was not necessarily the biggest observed effect among all models. The following set of hexagonally binned scatterplots in

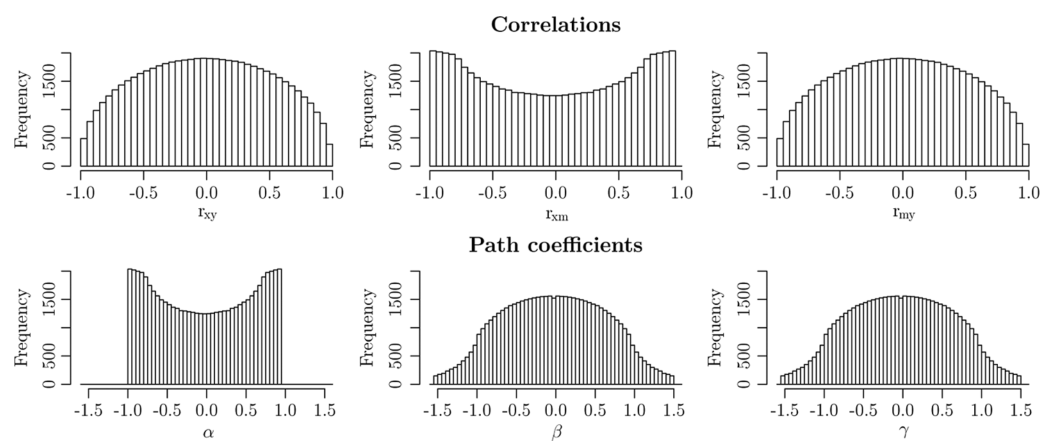

FIGURE 3 Histograms of correlations and path coefficients in the analytic study.

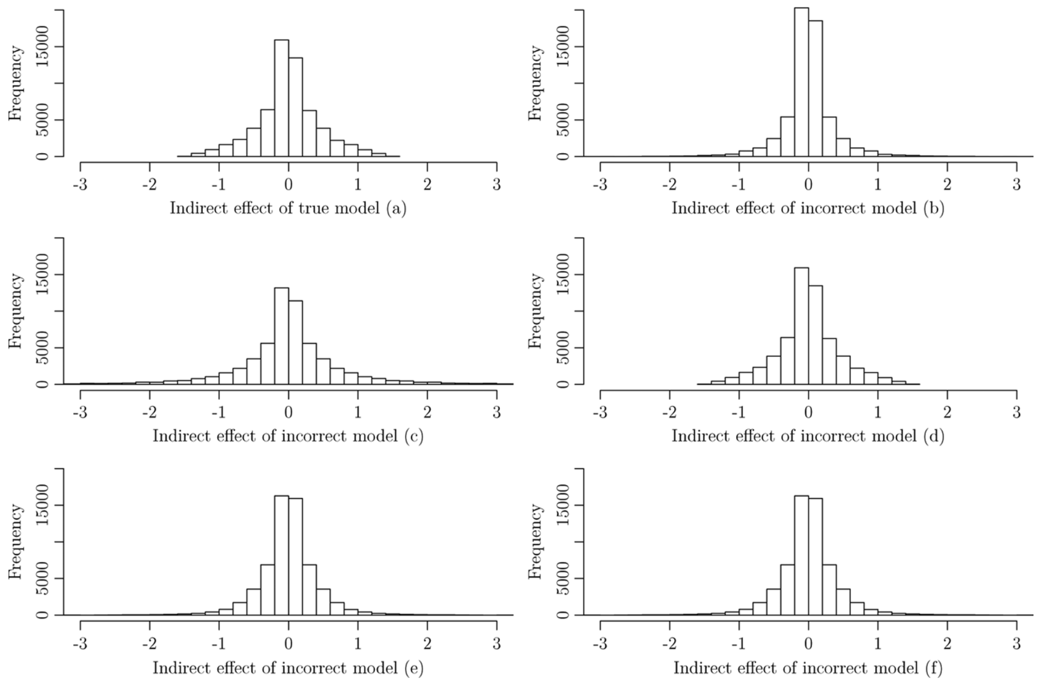

FIGURE 4 Histograms of indirect effects of all models from Figure 2 over the complete parameter space.

Figure 5 demonstrates that given a particular true indirect effect, it is quite possible that other models have larger indirect effects. As an example, consider the scatterplot that shows the size of the true indirect effect and the indirect effect that would be obtained when analyzing the data using the incorrect model in Figure 2d, in which X is incorrectly used as a mediator between M and Y. For every possible value of the true indirect effect, there are constellations of path coefficients that yield a potentially larger effect for the incorrect model. For example, in all constellations in which the true indirect effect is zero (a situation that emerges in all models in which either α, or β, or both are set to zero), the indirect effect of the incorrect model (although often being zero) can nevertheless range from around -1 to 1. This pattern holds for all incorrect models (b to f).

When comparing the indirect effects across all conditions and models, the true model yields the largest indirect effect in 18.7% of all cases. Given that there are six competing models, simply by chance we would expect to pick the correct model in 16.7% of all cases. This clearly demonstrates that declaring the mediation model with the largest indirect effect as the correct one is a strategy that will simply not work, because in the majority of cases, the applied researcher would select an incorrect mediation model.

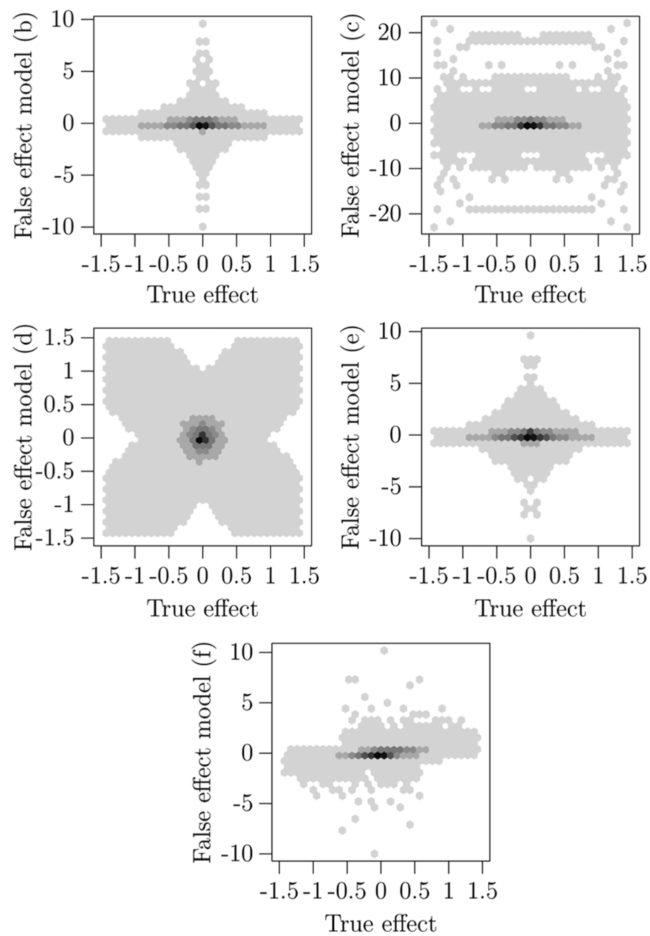

FIGURE 5 Binned scatterplots of the relationship between the true indirect effect and indirect effects derived from incorrect models in Figure 2.

Limitations

In our analytic study we did not assess the statistical significance of the observed effects but simply evaluated the analytically derived asymptotic magnitude of the coefficients. It is possible to derive standard errors of all regression coefficients analytically, given the true model and an assumed sample size *n*. With arbitrarily large sample sizes, all of the effects would, of course, attain significance, but even with smaller samples in which this would not be the case, we do not believe that our results would be impacted. The fact remains that alternative models in the equivalence class can have (often substantially) larger indirect effects and would therefore be chosen as the correct model, if an applied researcher would use the described decision rule.

A further possible objection to this finding is that we have considered a very large parameter space and that correlations of very large magnitude are very unlikely to be present in social science research. Perhaps the approach of picking the mediation model with the largest effect would fare better if one would only consider more realistic correlations (e.g., nothing that exceeds .3). As is already evident from the scatterplots in Figure 5, there will still be numerous situations in which incorrect models will be favored. Empirically, we can observe that if only models are considered that have (absolute) correlations smaller than $|.4|$, the correct model is chosen in 20.7% of all cases, with correlations smaller than $|.3|$, this increases to 23.0%, and with correlations smaller than $|.2|$, it increases again to 26.6%. Although those are slightly better results, they are still far from being close to be useful as a decision rule.

DISCUSSION

Reversing arrows in mediation models does not tell us whether one model is better than the other. The fact that one indirect effect shows up as being large, and another one as small, has no bearing on the veracity of the underlying model. The practice of testing alternative models that are in the same equivalence class should therefore be abandoned. Echoing Fiedler et al. (2011), we also assert that a mediator cannot be determined using statistics, especially if this includes the bad practice of checking different equivalent models as a basis for this decision. Even if a mediation model shows an indirect effect, all other equivalent models are equally endorsed, and thus mediation cannot be established by a (significant) indirect effect. The equivalent models can *only* be ruled out through design features, (e.g., randomization), or theoretical assumptions, (e.g., a priori knowledge that a variable cannot be influenced, e.g., biological gender). These assumptions are qualitative, causal assumptions and distinct from statistical assumptions, e.g., normality, or linearity.

Randomly assigning participants to, for example, X ensures that every equivalent model that posits that X is influenced by other variables can be ruled out—even if it would yield a larger indirect effect. Likewise, longitudinal designs can rule out alternative models. Again, as an example, if X was assessed prior to M, we can rule out models in which M has a direct effect on X (as this would indicate causation backward in time). For an overview of assumptions that one can make in the simple tri-variate model, along with a list of all equivalent models that are simultaneously endorsed, and which assumptions rule out which equivalent models, consult Thoemmes, Liao, and Yamasaki (2014).

The inability to use statistical evidence to distinguish between models in the same equivalence class is not limited to mediation models but applies to other models as well. Researchers should abandon attempts to argue for the veracity of one chosen model over another based on statistical evidence alone, if the two models are in the same equivalence class. Instead researchers must invoke and defend assumptions that lend support to their favored model. If one wants to make claims about directed (causal) effects, statistical and data-driven solutions are insufficient but theoretical assumptions are always necessary (Pearl, 2000).

In accordance with the title of the special issue on disadvantages of the mediation model, we observe that one such disadvantage is that the mediation model and the presence of an indirect effect are not actually a logical argument for the presence of a mediated causal process. As we have demonstrated analytically, equivalent models that are incorrect and are *not* mediation models can also yield large indirect effects, even though the assumed mediational process does not correspond to the true process.

ACKNOWLEDGEMENTS

The author would like to thank Nathan Spreng, Garett Howardson, and Sarah Moore for feedback, and Marina Yamasaki for derivation of some implied effects.

REFERENCES

Chen, B., & Pearl, J. (in press). Graphical tools for linear structural equation modeling. *Psychometrika*.

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, *47*, 1231–1236.

Fredrickson, B. L., Tugade, M. M., Waugh, C. E., & Larkin, G. R. (2003). What good are positive emotions in crisis? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *Journal of Personality and Social Psychology*, *84*, 365–376.

Hayduk, L., Cummings, G., Stratkotter, R., Nimmo, M., Grygoryev, K., Dosman, D., … Boadu, K. (2003). Pearl's d-separation: One more step into causal thinking. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*, 289–311.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 185–199.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Routledge.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*, 173–181.

Meek, C. (1995). Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (pp. 403–410). Morgan Kaufmann Publishers.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York, NY: Cambridge University Press.

Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research*, *21*, 309–331.

Thoemmes, F., Liao, W., & Yamasaki, M. (2014). *The Markov equivalence class of the mediation model*. Paper presented at the annual Modern Modeling Methods Conference, Storrs, CT.

Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*, *56*, 330–338.

APPENDIX A: SHORT INTRODUCTION TO D-SEPARATION

Given a graphical causal model (Pearl, 2000), we can test whether two variables are (conditionally) independent of each other, using the graphical criterion of d-separation. To test whether two variables are d-separated from each other, one enumerates all path tracings between these two variables (regardless of direction of arrows). If every path traced between the two variables contains a variable with two arrows pointing into it (referred to as a collider), then the two variables are d-separated, implying statistical independence. If there are paths that do not contain a collider (also referred to as unblocked paths), then the two variables are d-connected, implying statistical dependence.

To test whether two variables are conditionally independent from each other, using some conditioning set Z, we again enumerate all paths and then check for every path if it either contains a collider, or if any variable in the set of Z resides on the path, but is not a collider itself. In either of these two cases, the path becomes blocked, and if all enumerated paths are blocked, then the two variables are conditionally d-separated, implying conditional statistical independence. Note that if a variable in the set Z is a collider on a path, or the descendant of a collider (meaning that it has an arrow from a collider going into it), then this variable no longer blocks a path, leaving it potentially open, and thus implying that the two variables are not d-separated from each other.

In a classic tri-variate mediation model, none of the three variables are d-separated. We can easily verify this by considering the paths that connects each pair of variables in the mediation model. X and Y have two paths connecting them, $X \rightarrow M \rightarrow Y$, and $X \rightarrow Y$. Both paths are open, and after conditioning on M only the first path is blocked, but the second remains open, hence X and Y cannot be d-separated. Likewise X and M are connected via two paths, $X \rightarrow M$ and $X \rightarrow Y \leftarrow M$. The first path is open, and the second is closed due to collider Y. Conditioning on Y leaves the first path open and opens the previously closed second path, implying that the variables cannot be d-separated. Finally, M and Y, are connected by two paths, $M \rightarrow Y$ and $M \leftarrow X \rightarrow Y$. Both paths are open, and conditioning on X only closes the second path, again implying that the two variables cannot be d-separated. In summary, the classic tri-variate mediation model has no implied d-separations. Any other model with these three variables that also has no implied d-separation constraints is in the same equivalence class. As described in the article, all models that one could form by replacing any arrow of the mediation model with either another directed or bidirected arrow will also have no d-separation constraints, and will thus be in the same equivalence class.

APPENDIX B: DERIVATION OF INDIRECT EFFECTS USING PATH-TRACING RULES AND THE RECURSIVE FORMULA

Given the model in Figure 1a, the indirect effect is simply the product $\alpha\beta$. We may also derive implied correlations from the path coefficients using tracing rules (Wright, 1922) and then use the recursive formula to derive partial regression coefficients, and from those indirect effects. The use of standardized variables makes some of the computations easier, but the approach could be easily adapted to unstandardized variables. According to the tracing rules and assuming fully standardized variables, the following are the correlations of the three variables in the true model:

$$\begin{aligned}\rho_{XM} &= \alpha \\ \rho_{MY} &= \alpha + \beta\gamma' \\ \rho_{XY} &= \gamma' + \alpha\beta.\end{aligned}$$

All regression coefficients in a model can be expressed as a function of the correlations among the variables. Again assuming standardized variables, all simple bivariate regression coefficients can be expressed as simple bivariate correlations, for example, $\beta_{XY} = \beta_{YX} = \rho_{XY} = \rho_{YX}$. Using the recursive formula, we can also derive *partial* correlation and regression coefficients. We use a small dot to denote the variable that is partialed out of a relationship between variables that are preceded by the dot. When only one variable is partialled out, the partial correlation coefficient of variables X and Y, partialling for M, can be expressed as

$$\rho_{XY \cdot M} = \frac{\rho_{XY} - \rho_{XM}\rho_{MY}}{\sqrt{1-\rho_{XM}^2}\sqrt{1-\rho_{MY}^2}}.$$

From this formula, once can derive an expression for the partial standardized regression coefficient, using differential calculus (Chen & Pearl, in press). The general formula is

$$\beta_{XY \cdot M} = \frac{\beta_{XY} - \beta_{XM}\beta_{MY}}{(1-\beta_{XM}^2)}.$$

As described previously, the simple standardized regression coefficients can be expressed using correlation coefficients, which yields

$$\beta_{XY \cdot M} = \frac{\rho_{XY} - \rho_{XM}\rho_{MY}}{(1-\rho_{XM}^2)}.$$

This, in turn, allows any of the (partial) regression coefficients to be expressed using the path coefficients defined in the initial graph. Here in this example, this would be

$$\beta_{XY \cdot M} = \frac{\alpha\beta + \gamma' - (\alpha(\beta + \alpha\gamma'))}{1-\alpha^2} = \frac{\alpha\beta + \gamma' - \alpha\beta - \alpha^2\gamma'}{1-\alpha^2} = \gamma'.$$

Note that for regression coefficients that correspond to the true graph, this exercise is needlessly complicated, and we could read the regression coefficients directly off the graph; however, the described procedure becomes useful when we want to generate regression coefficients that are obtained when an incorrect model is fitted to a true data generating model.

In summary, we now have a way to first (a) derive correlation coefficients from a true path model using tracing rules, then (b) derive all (partial) regression coefficients (and indirect effects from those) expressed as simple regression coefficients, and (c) express all (partial) regression coefficients as a function of the path coefficients in the graph.

The following are standardized (partial) regression coefficients for the true model in our examples:

$$\begin{aligned}
\beta_{XM} &= \rho_{XM} = \alpha \\
\beta_{MY \cdot X} &= \frac{\rho_{MY} - \rho_{XM}\rho_{XY}}{(1-\rho_{XM}^2)} = \beta \\
\beta_{XY \cdot M} &= \frac{\rho_{XY} - \rho_{XM}\rho_{MY}}{(1-\rho_{XM}^2)} = \gamma'.
\end{aligned}$$

The indirect effect is the product of $\beta_{MY \cdot X}$ and β_{XM}, which is simply $\alpha\beta$. We can also use the same approach to derive path coefficients for the incorrect model in Figure 1 (b). They are:

$$\begin{aligned}
\beta_{XY} &= \rho_{XY} = \alpha\beta + \gamma' \\
\beta_{YM \cdot X} &= \frac{\rho_{MY} - \rho_{XM}\rho_{XY}}{(1-\rho_{XY}^2)} = \frac{\beta - \alpha^2\beta}{1-(\alpha\beta+\gamma')^2} \\
\beta_{XM \cdot Y} &= \frac{\rho_{XM} - \rho_{MY}\rho_{XY}}{(1-\rho_{MY}^2)} = \frac{\alpha - ((\alpha\beta+\gamma')(\beta+\alpha\gamma'))}{1-(\alpha\beta+\gamma')^2}.
\end{aligned}$$

The indirect effect in the wrong model would be the product of β_{XY} and $\beta_{XM \cdot Y}$.

Using this method we can derive all possible (partial) regression coefficients that could be formed based on the true model in Figure 2a and express them using path coefficients:

$$\begin{aligned}
\beta_{XY} &= \beta_{YX} = \alpha\beta + \gamma' \\
\beta_{XM} &= \beta_{MX} = \alpha \\
\beta_{MY} &= \beta_{YM} = \beta + \alpha\gamma' \\
\beta_{XY \cdot M} &= \gamma' \\
\beta_{XM \cdot Y} &= \frac{(\alpha - ((\alpha\beta+\gamma')(\beta+\alpha\gamma')))}{(1-(\alpha\beta+\gamma')^2)} \\
\beta_{MY \cdot X} &= \beta \\
\beta_{YX \cdot M} &= \frac{(\alpha^2\gamma' + \gamma')}{(1-(\alpha\gamma'+\beta)^2)} \\
\beta_{MX \cdot Y} &= \frac{(\alpha - (\alpha\beta^2 + \beta\gamma' + \alpha^2\beta\gamma' + \alpha\gamma'^2))}{(1-(\beta+\alpha\gamma')^2)} \\
\beta_{YM \cdot X} &= \frac{(\beta - \alpha^2\beta)}{(1-(\alpha\beta+\gamma')^2)}.
\end{aligned}$$

These are the exact same formulas that were used for the analytic study that examined all possible indirect effects of the models in Figure 2.

APPENDIX C: REPLICATION R CODE

```
library(matrixcalc)
#True mediation model X-M-Y
alpha <- seq(-1.5,1.5,.05)
beta <- seq(-1.5,1.5,.05)
gammap <- seq(-1.5,1.5,.05)

#expand grid to have all possible combinations
grid1 <- expand.grid(alpha,beta,gammap)
names(grid1) <- c("a","b","gp")

#generate correlations from values in grid
grid1$rxm <- grid1$a
grid1$rxy <- grid1$gp + grid1$a*grid1$b
grid1$rmy <- grid1$b +  grid1$a*grid1$gp

#check that resulting correlation matrix is positive defininte
res <- matrix()
for (i in 1:nrow(grid1))
{
  t <- is.positive.definite(matrix(c(1,grid1$rxm[i],
  grid1$rxy[i],grid1$rxm[i],1,grid1$rmy[i],grid1$rxy[i],
  grid1$rmy[i],1),nrow=3,byrow=TRUE))
  res[i] <- t
}
grid1$check <- res
grid2 <- grid1[grid1$check==TRUE,]

#form and rename final grid and clean up workspace
fgrid <- grid2
rm(grid1,grid2,i,res,t)

#now add indirect effects
#using the following nomenclature
#trueind = true effect from x-m-y model
#falseind1 = false effect from x-y-m model
#falseind2 = false effect from m-y-x model
#falseind3 = false effect from m-x-y model
#falseind4 = false effect from y-x-m model
#falseinf5 = false effect from y-m-x model

#attach fgrid to make formulas easier to digest
attach(fgrid)
```

```
#regression coefficients
fgrid$betaxy <- a*b+gp
fgrid$betayx <- a*b+gp
fgrid$betaxm <- a
fgrid$betamx <- a
fgrid$betamy <- b + a*gp
fgrid$betaym <- b + a*gp

#partial regression coefficients
fgrid$betaxy.m <- gp
fgrid$betaxm.y <- (a-((a*b+gp)*(b+a*gp))) / (1-(a*b+gp)^2)
fgrid$betamy.x <- b
fgrid$betayx.m <- (a^2*gp+gp) / (1-(a*gp+b)^2)
fgrid$betamx.y <- (a-(a*b^2+b*gp+a^2*b*gp+a*gp^2)) / (1-(b+a*gp)^2)
fgrid$betaym.x <- (b-a^2*b) / (1-(a*b+gp)^2)

#indirect effects as functions of partial regression coefficients
fgrid$trueind <-   fgrid$betaxm * fgrid$betamy.x
fgrid$falseind1 <- fgrid$betaxy * fgrid$betaym.x
fgrid$falseind2 <- fgrid$betamy * fgrid$betayx.m
fgrid$falseind3 <- fgrid$betamx * fgrid$betaxy.m
fgrid$falseind4 <- fgrid$betayx * fgrid$betaxm.y
fgrid$falseind5 <- fgrid$betaym * fgrid$betamx.y

#attach grid again to include new variables
attach(fgrid)

#get the absolute max of the columns trueind to falseind5

fgrid$max <- apply(abs(fgrid[,c(20:25)]),1,max)

#get a binary decision indicator
#1 means correct, 0 means incorrect
fgrid$decision <- ifelse(fgrid$max==abs(fgrid$true),1,0)

#table with counts of correct vs incorrect
table(fgrid$decision)
```