THIS IS A **PREPRINT** AND HAS NOT BEEN PEER-REVIEWED

The Indirect Effect is Omitted Variable Bias. A Cautionary Note on the Theoretical
Interpretation of Products-of-coefficients in Mediation Analyses.

Lennert Coenen

Tilburg Center for Cognition and Communication, Tilburg University, Tilburg
University, The Netherlands;

Leuven School for Mass Communication Research, Faculty of Social Sciences, KU
Leuven, Leuven, Belgium;

Research Foundation – Flanders (FWO)

Author note

This work was partly supported by grant 12J7619N from Research Foundation –
Flanders (FWO). Correspondence concerning the preprint should be addressed to
Lennert Coenen, Tilburg Center for Cognition and Communication, School of
Humanities and Digital Sciences, Tilburg University, PO Box 90153, 5000 LE Tilburg,
The Netherlands. Tel: +31 (0)13 466 8118. E-mail: l.coenen@tilburguniversity.edu.

The Indirect Effect is Omitted Variable Bias. A Cautionary Note on the Theoretical Interpretation of Products-of-coefficients in Mediation Analyses.

This paper intends to remind communication scientists that the indirect effect as estimated in mediation analyses is a statistical synonym for omitted variable bias (i.e., confounding or suppression). This simple fact questions the interpretability of statistically significant ‘indirect effects’ when using observational data: in social reality all variables correlate with each other to some extent - the so-called ‘crud factor’ - which means that omitted variable bias and ‘indirect effects’ at the population level are virtually guaranteed regardless of the actual variables involved in the statistical mediation model. As a result, there can be no inferential link between the observation of a significant indirect effect and a theoretical claim of mediation. Through this argument the paper hopes to add to the existing warnings on mediation analyses and cultivate a more critical interpretation of ‘indirect effects’ in communication science.

Keywords: mediation analysis; indirect effect; omitted variable bias; statistical inference; significance

Establishing the working mechanisms and explanatory processes underlying communication phenomena is considered to be one of the most important goals in communication science (e.g., Holbert & Stephenson, 2003; Preacher & Hayes, 2008; Valkenburg, Peter, & Walther, 2016). To contribute to this end researchers often resort to a statistical technique known as mediation analysis: a statistical mediation model defines the relationship between three sets of variables - a set of explanatory variables, consisting of (presumed) causal predictors and covariates, a set of mediators, and a set of dependent variables. In the simplest case this boils down to a model consisting of one observed independent variable x , one observed mediator m , and one observed dependent variable y (see Figure 1). The underlying theoretical hypothesis is that variable x *indirectly* influences y through m or, in other words, that variable m serves as the *mediator* of the effect of x on y . The associated statistical test proceeds by estimating

the product of regression coefficients (b_1b_3) from two separate linear regression models:

$$m_i = b_{01} + b_1x_i + e_{i1}, \quad (1)$$

where b_1 represents the effect of x on m , and

$$y_i = b_{02} + b_2x_i + b_3m_i + e_{i2}, \quad (2)$$

where b_3 represents the effect of m on y , controlling for the independent variable x .

Statistical inference on the indirect effect b_1b_3 typically is usually done by evaluating the null hypothesis that $\beta_1\beta_3 = 0$, which can be achieved either by calculating a standard error and test statistic for b_1b_3 under the assumption that it has an asymptotically normal sampling distribution (Sobel, 1987), or by bootstrapping the empirical sampling distribution of b_1b_3 , its standard error, and confidence interval (Shrout & Bolger, 2002). Modern software packages such as PROCESS (Hayes, 2018) and RMediation (Tofighi & MacKinnon, 2011) have made these inferential procedures easily accessible (even for relatively complicated multiple mediator models), so it is hardly surprising to see mediation analyses being so widely used in the communication literature. To put this into perspective, Chan, Hu, & Mak (2020) recently showed that the share of communication studies relying on mediation analyses has been on a quick rise: in 2007, only 7% of articles in the communication literature reported mediation results; by 2017, this figure had already increased to 22%.

FIGURE 1 ABOUT HERE

A well-known problem: significant indirect effects do not validate causal interpretations.

Despite the popularity of mediation analyses, methodologists have long warned against their application, especially in research using non-experimental data (e.g., Bullock, Green, & Ha, 2010; Bullock & Ha, 2011; Kline, 2015). The reason for this is that the results of mediation analyses have little to say about the implied causality underlying a mediation hypothesis: the mediation analysis itself does nothing to establish temporal order or rule out third-variable explanations and, as such, significant indirect effects in observational studies should not be taken as direct evidence of a specific causal pathway (Fiedler, Schott, & Meiser, 2011; Kline, 2015). Of course, the basic distinction between causality and correlation is thoroughly acknowledged in limitation sections of communication research, and few researchers would *literally* claim to have found definitive evidence for a causal mechanism based on a significant ‘indirect effect’ alone. However, Chan and colleagues (2020) recently found that communication scholars do, in fact, use causal terminology to describe significant indirect effects, suggesting that mediation analyses are at least implicitly assumed to provide evidence about process-oriented hypotheses. This is also attested to by the fact that the widespread adoption of mediation has been explicitly applauded by communication scholars – as a sign of theoretical progress and methodological sophistication in the field (e.g., Perloff, 2013).

Unfortunately, there is good reason to argue that significant indirect effects often do not provide any evidence - not even tentative evidence - for process-related theoretical claims. This is not just true because there might be omitted confounders or alternative variable orders at play – an issue that has been acknowledged and discussed at length (Chan et al., 2020; Fiedler et al., 2011; Kline, 2015). Arguably, an even more fundamental issue is that, in many cases, the inferential test of the indirect effect is

nothing short of a logical truism. This is also the point stressed throughout this paper: it is statistically guaranteed that any random constellation of three observed variables will generate a significant ‘indirect effect’ in a mediation model at some fixed but unknown sample size n . This is true regardless of whether or not the variables in the mediation model are actually related in any theoretically meaningful way, which means that there is little value at all in interpreting ‘significant’ products-of-coefficients from mediation analyses as meaningful evidence of some sort.

A lesser-known problem: significant indirect effects are logical truisms.

To understand this criticism it is enough to reiterate the basic (though rarely considered) fact that mediation is a statistical synonym for confounding or suppression (see also MacKinnon, Krull, & Lockwood, 2000). This is easily seen when y is written as a function of x alone. Given (1) and (2) it holds that:

$$\begin{aligned} y_i &= b_{02} + b_2 x_i + b_3(b_{01} + b_1 x_i + e_{i1}) + e_{i2}, \\ &= (b_{02} + b_3 b_{01}) + (b_2 + b_1 b_3) x_i + (b_3 e_{i1} + e_{i2}), \end{aligned} \quad (3)$$

which can be rewritten as

$$y_i = b_{03} + b_4 x_i + e_{i4}, \quad (4)$$

such that

$$b_4 = b_2 + b_1 b_3. \quad (5)$$

In the terminology of mediation analysis Equation 5 is typically said to reflect that the Total Effect b_4 equals the sum of the Direct (b_2) and Indirect Effects ($b_1 b_3$) of x on y . This choice of words is somewhat unfortunate, however, as it immediately attaches a causal connotation to regression coefficients and risks muddling the literal statistical

meaning of Equation 5: literally, Equation 5 just shows that the overall, bivariate linear relationship between x and y is equal to the sum of the relationship *after controlling* for m and a residual term. The residual term – called ‘the indirect effect’ in mediation terminology – thus simply represents the difference in a regression coefficient before and after controlling for a variable m :

$$b_1b_3 = b_4 - b_2. \quad (6)$$

When written in this form, the residual term b_1b_3 is also referred to as *omitted variable bias* in the statistical literature (e.g., Cinelli & Hazlett, 2020): it is the under- or overestimation in b_4 if we fail to control for m in the prediction of y in Equation 4.

The fact that omitted variable bias and indirect effects are mathematical synonyms has two logical implications. First, it means that there will always be omitted variable bias in a population-level model if m serves as a mediator of the x - y -relationship in the population. This makes conceptual sense, because for a variable to be a mediator it needs to explain shared variance between x and y . This reasoning was also made explicit in traditional mediation techniques such as Judd and Kenny’s (1981) causal-steps approach, where a change in the relationship between x and y after controlling for variable m was taken as evidence for statistical mediation. A second but less intuitive – and, arguably, less appealing – implication of Equation 5 is that any variable m acting as a population-level confounder or suppressor of the x - y -relationship will serve as a statistical mediator in the population-level statistical model. Indeed, the population-level counterpart of Equation 6 implies that there will be no mediation or omitted variable bias at the population level *if and only if* the product of population-level regression coefficients $\beta_1\beta_3 = 0$. In any other case a type of both can be inferred: when $\beta_1\beta_3\beta_2 < 0$ there will be statistical suppression and inconsistent mediation; when

$\beta_1\beta_3\beta_2 > 0$ one may infer statistical confounding and consistent mediation. This means that any variable m somehow related to both x and y in the population is guaranteed to generate a population-level ‘indirect effect’ $\beta_1\beta_3 \neq 0$ in the statistical model.

While this fact might not come as a surprise to statistically versed readers, it seems important to reiterate as it underlines that, in many common cases, testing for indirect effects in communication research is nigh meaningless. In essence, the statistical test of the ‘indirect effect’ only serves to address the question: does variable m serve as an omitted variable – a confounder or suppressor - in the relationship between x and y ? When posited like this, a test of statistical mediation might not appear all that interesting anymore, and it is even less so when we consider that the answer will nearly always be “yes” when a study relies on observational data. The reason for this is known as the *crud factor*: “in the social sciences, everything is somewhat correlated with everything” (Meehl, 1990, p. 108). All social, behavioral, and personality variables are part of a complex, intractable constellation of interdependencies, which means that any randomly measured set of such variables can be expected to be at least *somewhat* interrelated at the population level (see also Orben & Lakens, 2020). But if it is true that virtually all relationships between observed variables at the population level are nonzero, then it follows that virtually no omitted variable bias in a finite linear constellation of variables will ever be literally zero. And if virtually no omitted variable bias in a finite linear constellation of variables will ever be literally zero, then, by Equation 6, no so-called ‘indirect effect’ in a population-level statistical model will ever be equal to zero! This means that any ‘indirect effects’ coefficient b_1b_3 is guaranteed to reach statistical significance given an unknown but fixed sample size n , regardless of the theoretical relationships of variables that were involved in the mediation model to begin with. In other words, the statistical evidence that is typically used to evaluate

mediation hypotheses in observational communication research does not deliver any meaningful evidence at all. It rather represents what Mayo (2018) calls BENT evidence (Bad Evidence No Test): even if the underlying process-oriented theory were *not* true, a large enough sample size would guarantee the significance of ‘the indirect effect’, meaning that there is no inferential link whatsoever between the statistical result and the underlying theoretical hypothesis.

Again, it seems useful to underline that this criticism is different in spirit from more oft-heard cautions against mediation analyses in observational designs: the problem addressed here is *not* with the fact that b_1b_3 might be biased because of *other* theoretically plausible explanations (i.e., other omitted variables; Fiedler et al., 2011; Pieters, 2017). The point is that indirect effect b_1b_3 is *itself* a reflection of m being an omitted variable in Equation 4, which – per the crud factor – is virtually guaranteed and therefore not scientifically meaningful. The semantic distinction between these two arguments becomes clear when we consider the research question that guided Mak and colleagues’ (2020) critique of mediation results: the authors wanted to know “how confident [we can] be that the patterns of observations among the independent variable(s), mediator(s), and dependent variable(s) (i.e., the mediation model) reported in studies *can be inferred to the real world*”? (Mak et al., 2020, emphasis added). From the perspective of their analysis, the inference is rarely warranted because of the lacking design choices in tests of mediation. From the perspective of the current paper, however, the answer to the question is very different (albeit equally critical): if by ‘pattern of observations’ we refer to the rejection of a null hypothesis for the indirect effects coefficient – which is generally the focal point of published mediation analyses - we will almost certainly be able to infer that pattern to the real world! That is, non-null values for $\beta_1\beta_3$ are guaranteed because of the crud factor, so rejecting the null

hypothesis for b_1b_3 will almost certainly be the correct decision. However, the fact that inferring $\beta_1\beta_3 \neq 0$ is nearly always warranted, and the fact that we know this purely on logical grounds, suggests that a statistical test of mediation itself is inherently meaningless and not worth conducting!

Does crud really render observational indirect effects meaningless?

While the crud-criticism fundamentally challenges the value of testing for indirect effects one could formulate several rebuttals to it. First, not all readers will be convinced that the crud factor plays such a fundamental role in communication research. It has been argued, for instance, that the crud factor is an empirical hypothesis in itself, and that there is “no a priori reason to believe that one will always reject the null hypothesis at any given sample size” (Mulaik, Raju, & Harshman, 1997, p. 80). One problem with this counterargument is that it is not feasible to ask for the crud factor to be tested (as it would require us to exhaust population-level observations, which is impossible). Another problem is that the assumption of there being a crud factor seems much more parsimonious than the assumption of there being no such thing: saying that a crud factor does *not* exist requires us to take seriously that exact, literal null parameters occur in social reality (and, in fact, that they are prevalent). For a correlation coefficient this would imply that exhausting the population of possible observations would have us end up with a correlation of $\rho = 0$ – that is, not $\rho = 0.001$ or $\rho = 10^{-10}$, but *exactly* $\rho = 0$. This seems highly unlikely, especially when we consider that what we consider as a population parameter in social science does not actually exist as a fixed constant but depends on the exact moment at which the population is investigated and statistically defined. People are not invariant or interchangeable, which means that there is a range of possible values for a parameter depending on the exact moment of sampling (a hyperdistribution of sort). And if a given population parameter is

conceptualized as a random draw from a hyperdistribution of possible parameters, exact null values at the moment of sampling will become increasingly unlikely when the number of parameters in the hyperdistribution increases. Put differently, as long as we do not want to assume that naturally occurring parameters in social reality are somehow predetermined to be exactly 0, it is highly unlikely of there being relationships in social reality that are exactly equal to 0. The crud factor is generally considered as an axiom by statisticians, and it has been used as a key warning against the widespread reliance on null hypothesis testing. As Cohen (1990, p. 1308) noted:

A little thought reveals a fact widely understood among statisticians: The null hypothesis, taken literally (and that's the only way you can take it in formal hypothesis testing), is always false in the real world. It can only be true in the bowels of a computer processor running a Monte Carlo study (and even then a stray electron may make it false). If it is false, even to a tiny degree, it must be the case that a large enough sample will produce a significant result and lead to its rejection.

But if this is true, then it necessarily follows that the inference of non-zero omitted variable bias (through the statistical rejection of a non-null 'indirect effect') is a truism.

Another rebuttal to the paper could be that the crud factor argument is actually a criticism of null-hypothesis testing in general, rather than a criticism of mediation analyses per se. That is certainly correct, but the crud factor poses problems that are particularly pressing in the context of mediation analyses. There are at least three reasons for this. First, as mentioned earlier, the mediation literature relies very heavily on causal terminology (indirect, direct, and total 'effects') that is typically avoided when describing correlational findings in observational research. Such a use of terminology is reasonable as long as one sticks to the theoretical rationale guiding research efforts, but it becomes problematic when the rejection of a null value for a

product of coefficients is also taken as an empirical corroborator for an ‘indirect effect’.

With this in mind it seems valuable to always entertain the literal interpretation of the product of coefficients as omitted variable bias: instead of interpreting b_1b_3 as the ‘indirect effect of x on y through m ’ one may simply think of it as ‘the difference in the relationship between x and y before and after controlling for m ’. In models with more than one mediator this reasoning can be extended: what is typically called the ‘Total Indirect Effect’ can be rephrased as ‘the difference in the relationship between x and y before and after controlling for *all* m ’, and the specific ‘indirect effects’ are contributions of specific m ’s to the total omitted variable bias. The value of this formulation is that it carries no causal connotation, thus clarifying that researchers need to put considerable effort into explaining why some non-null product of coefficients is a corroborator for a process-related theoretical claim rather than tautological evidence for omitted variable bias.

A second reason why the crud factor is pressing in the context of mediation analyses is that there is no commonly agreed-upon metric of effect size to evaluate the strength of an ‘indirect effect’. While various different metrics have been proposed, many of them are known to have statistical and conceptual issues, and a consensus on their application is yet to be found (see Lachowicz, Preacher, & Kelley, 2018, for a discussion and promising development). This is unfortunate, because an estimate of effect size may help counter allegations of crud-factor relationships: crud-factor relations, being the outcome of a complex chain of interdependencies, are likely to be very small in size. For this reason, a relatively strong effect size already lends more evidence for some theoretically viable indirect effect rather than crud per se. This is why it seems crucial for researchers to substantively interpret at least *some* type of available effect size – the most accessible simply being the estimated products-of-

coefficients in unstandardized or standardized form (the latter is known as the index of mediation: Preacher & Hayes, 2008). As noted above, the interpretation should focus on the reasons why the size of these coefficients can be thought to reflect corroborations of the indirect effects hypotheses rather than just crud-like omitted variable bias. This issue currently not properly addressed in the communication literature; mediation in the field is simply inferred through the statistical significance of products-of-coefficients (Chan et al., 2020).

A third reason why the crud factor problem is pressing in the context of mediation analysis is that establishing statistical mediators is commonly believed to be a more impressive theoretical contribution than simply finding *x-y*-relationships (Hayes, Preacher, & Myers, 2011). It should be clear by now that this is not necessarily the case: the crud factor guarantees non-null ‘indirect effects’ in observational studies, so there is no principled reason to value research more if it reports some significant indirect effects. On the contrary, there is a fundamental risk for a literature relying too heavily on these types of findings: if mediation hypotheses passing null hypothesis tests are considered to be theoretical contributions in a field, then nearly all process-related conjectures - regardless of their validity or logical content – may be counted as theoretical contributions. As a result, the knowledge base of the discipline risks diverging into a patchwork of idiosyncratic ‘process models’ that are not necessarily meaningful and have no clear connection to one another (see also Rohrer, Rühnermund, Arslan, & Elson, 2020). This clarifies an additional benefit of adopting a more critical attitude toward mediation analyses: if we no longer consider significant indirect effects as meaningful corroborators for process-related hypotheses, then we might prevent communication theory from further increasing “in complexity without increasing in explanatory power” (Lang, 2013, p. 14).

Conclusion: the scientific insignificance of significant indirect effects

In sum, the takeaway message of this paper is pretty straightforward: the communication literature should be much more critical toward the theoretical viability of significant ‘indirect effects’. While this recommendation is certainly not new, it deserves being reiterated - especially given that previous cautions do not appear to have changed scholarly practice all that much (as is reflected in Chan et al., 2020). The arguments raised in the current paper also used a different frame compared to previous discussions: rather than reiterating that mediation results can be plagued by third variables or reverse causal order, this paper unpacked a simple statistical argument to show that the significance of indirect effects is nothing but a truism. Hopefully, this alternative perspective can convince communication scholars that null rejections alone have very little to say about the theoretical viability of an indirect effect.

The only circumstance under which null hypothesis tests for indirect effects are valuable is when a study uses an experimental design with random assignment for both the independent variable and the mediator (see Bullock et al., 2010; Bullock & Ha, 2011): if subjects are randomly assigned to conditions for x and m , the crud factor is no longer a concern because randomization and manipulation break pre-existing dependencies between all variables in the mediation model. However, the same is not true when only the independent variable x is manipulated! While such a set-up breaks the crud-factor dependency between x and m , and between x and y , the relationship between m and y remains untouched. This means that as long as the manipulation of x has a non-zero population-level influence on m , a population-level indirect effect on y is again guaranteed. Under conditions when it is not feasible to manipulate both x and m – for instance, when a research question requires naturalistic observation – mediation analyses might still have their place. However, under such circumstances, researchers will need to provide a very sound theoretical and methodological argument to embed

their interpretations – not just when it comes to the order of variables and their accounting for confounders (to address inverse causality and third variable explanations), but also when it comes to the size of the indirect effect (to address crud). An Editorial requirement to always report Directed Acyclic Graphs together with the sizes of the respective relationships might certainly help in this regard (Pearl, 2009; Rohrer, 2018). In any case, the statistical significance of a product-of-coefficients alone should no longer be considered as a meaningful basis to evaluate process-related claims in communication science.

References

- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, 98(4), 550–558. doi:10.1037/a0018933
- Bullock, J., & Ha, S. E. (2011). Mediation Analysis is Harder Than it Looks. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science* (pp. 508–521). New York, NY: Cambridge University Press.
- Chan, M., Hu, P., & K. F. Mak, M. (2020). Mediation Analysis and Warranted Inferences in Media and Communication Research: Examining Research Design in Communication Journals From 1996 to 2017. *Journalism & Mass Communication Quarterly*. doi:10.1177/1077699020961519
- Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1), 39-67.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45(12), 1304–1312. doi:10.1037/0003-066x.45.12.1304

- Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231–1236.
doi:10.1016/j.jesp.2011.05.007
- Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (2nd Edition). NY: Guilford Press.
- Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. In E. P. Bucy & R. Lance Holbert (Eds.), *Sourcebook for political communication research: Methods, measures, and analytical techniques* (pp. 434 – 465). New York, NY: Routledge
- Holbert, R. L., & Stephenson, M. T. (2003). The Importance of Indirect Effects in Media Effects Research: Testing for Mediation in Structural Equation Modeling. *Journal of Broadcasting & Electronic Media*, 47(4), 556–572.
doi:10.1207/s15506878jobem4704_5
- Judd, C. M., & Kenny, D. A. (1981). Process Analysis. *Evaluation Review*, 5(5), 602–619. doi:10.1177/0193841x8100500502
- Kline, R. B. (2015). The Mediation Myth. *Basic and Applied Social Psychology*, 37(4), 202–213. doi:10.1080/01973533.2015.1049349
- Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, 23(2), 244–261.
doi:10.1037/met0000165
- Lang, A. (2013). Discipline in Crisis? The Shifting Paradigm of Mass Communication Research. *Communication Theory*, 23(1), 10–24. doi:10.1111/comt.12000
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, 1, 173-181.

- Mayo, D. (2018). *Statistical Inference as Severe Testing. How to Get Beyond the Statistics Wars*. Cambridge: Cambridge University Press.
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1(2), 108-141.
- Mulaik, S. A., Raju, N. S., & Harshman, R. A. (1997). There is a time and a place for significance testing. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 65–115). Mahwah, NJ: Lawrence Erlbaum Associates, Inc
- Orben, A., & Lakens, D. (2020). Crud (re) defined. *Advances in Methods and Practices in Psychological Science*, 3(2), 238-247.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Perloff, R. M. (2013). Progress, paradigms, and a discipline engaged: A response to Lang and reflections on media effects research. *Communication Theory*, 23(4), 317-333.
- Pieters, R. (2017). Meaningful Mediation Analysis: Plausible Causal Inference and Informative Communication. *Journal of Consumer Research*, 44(3), 692–716. doi:10.1093/jcr/ucx081
- Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Thousand Oaks, CA: Sage.
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1). 27-42. doi: 10.1177/2515245917745629.

- Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2021). That's a lot to PROCESS! Pitfalls of Popular Path Models. PsyArXiv.
<https://doi.org/10.31234/osf.io/paeb7>
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422–445. doi:10.1037/1082-989x.7.4.422
- Sobel, M. E. (1987). Direct and Indirect Effects in Linear Structural Equation Models. *Sociological Methods & Research*, 16(1), 155–176.
doi:10.1177/0049124187016001006
- Tofighi, D., & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods*, 43(3), 692–700.
doi:10.3758/s13428-011-0076-x
- Valkenburg, P. M., Peter, J., & Walther, J. B. (2016). Media effects: Theory and Research. *Annual Review of Psychology*, 67. doi: 10.1146/annurev-psych-122414-033608.

Figure 1. A simple three-variable mediation model.