DRAFT 27 December 2024

Errors-in-Variables Regression as a Viable Approach to Mediation Analysis with Random Error-Tainted Measurements: Estimation, Effectiveness, and an Easy-to-Use Implementation

Andrew F. Hayes

University of Calgary

Paul D. Allison

Statistical Horizons LLC and University of Pennsylvania

Sean M. Alexander

University of Calgary

NOTE: A version of this manuscript is in the process of being peer reviewed. Please consult the published version once available

Corresponding author

Andrew F. Hayes
Haskayne School of Business
University of Calgary
2500 University Drive NW
Calgary, AB, Canada, T2N 1N4
andrew.hayes@ucalgary.ca

Abstract

Mediation analysis, popular in many disciplines that rely on behavioral science data analysis techniques, is often conducted using ordinary least squares (OLS) regression analysis methods. Given that one of OLS regression's greatest weaknesses is its susceptibility to estimation bias that results from unaccounted-for random measurement error in variables on the right-hand sides of the equation, many published mediation analyses certainly contain some and perhaps substantial bias in the direct, indirect, and total effects. In this manuscript, we offer errors-in-variables (EIV) regression as an easy-to-use alternative to OLS regression when a researcher has reasonable estimates of the reliability of the variables in the analysis. In three real-data examples, we show that EIV regression-based mediation analysis produces estimates that are largely equivalent to those obtained using an alternative, more analytically complex approach that accounts for measurement error—single-indicator latent variable structural equation modeling—yet quite different from the results generated by standard OLS regression that ignores random measurement error. In a small-scale simulation, we also establish that EIV regression successfully recovers the parameters of a mediation model involving variables adulterated by random measurement error while OLS regression generates biased estimates. To facilitate the adoption of EIV regression, we describe an implementation in the PROCESS macro for SPSS, SAS, and R that we believe now eliminates most any excuse one can conjure for not accounting for random measurement error when conducting a mediation analysis.

Errors-in-Variables Regression as a Viable Approach to Mediation Analysis with Random Error-Prone Measurements: Estimation, Effectiveness, and an Easy-to-Use Implementation

Researchers throughout the world, be they in the behavioral or natural sciences, business and management, public health, medicine, and many other fields, are interested in establishing through the scientific method those things that produce causal effects of relevance to their theories and their discipline and its application. But the goals of research often go beyond merely establishing that such effects exist. Researchers also seek to understand the underlying mechanisms by which those effects operate. *Mediation* is the term often employed when discussing the mechanisms that transmit causal effects, and researchers regularly theorize about and test *mediation models* using *mediation analysis*.

An example of a mediation analysis in action can be found in Grisbook, Dewey, and Cuthbert et al. (2024), who estimated a model examining posttraumatic stress and post-partum depression as mediators of the effect of emergency Caesarean section relative to spontaneous vaginal delivery and planned C-section on internalizing and externalizing behaviors of the child two years after birth. They found that women who delivered through emergency C-section, relative to planned C-section or vaginal birth, reported greater posttraumatic stress and post-partum depression three months later, which in turn was related to greater internalizing and externalizing behaviors in the child two to three years later. These results suggest that the form of delivery a woman experiences during childbirth can affect the child's behavior years later as a result, at least in part, of the psychological experiences and state of the mother that result from different delivery methods.

Additional examples of mediation analysis are abundant in the research literature (e.g., Gaboury, Belleville, & Lebel et al., 2023; Neufeld & Malin, 2022; Rokeach & Wiener, 2022; Smith, Andruski, Deng, & Burnham, 2022; Yasuda & Goegen, 2023). Indeed, when browsing any issue of journals that publish empirical research, it is common to find at least one mediation analysis in its pages (e.g., Chan, Hu, & Mak, 2022; Hayes & Scharkow, 2013; Pieters, 2017). And methodology articles about mediation analysis are among some of the most highly cited papers in the journals they are published in and in behavioral sciences methodology as a whole. This reflects the widespread popularity and transdisciplinary relevance of mediation analysis.

Mediation analysis is typically conducted using some kind of linear model-based path analysis, such as a set of ordinary least squares (OLS) regression analyses or a simultaneous estimation system such as observed variable structural equation modeling (SEM). Well-known among methodologists but not everyone who applies the work of methodologists is the deleterious effects of random measurement error in the variables being analyzed on the estimation of the effects of those variables in linear models (Bollen, 1989; Buonaccorsi, 2010; Cohen, Cohen, West, & Aiken, 2003; Darlington & Hayes, 2017; Shear & Zumbo, 2013). Random measurement error is ubiquitous in research that behavioral scientists conduct and is difficult to avoid entirely when measuring constructs that researchers study. But in practice, random measurement error is often ignored by researchers at the analysis phase, and this can produce bias in the estimation of those effects and can invalidate inferential tests of the effects that come out of an analysis, including a mediation analysis (e.g., Cole & Preacher, 2014). Advice offered by methodologists to counteract the effects of random measurement error include minimizing it at the design phase by using good measurement instruments, using SEM with a measurement model that captures random measurement error, or utilizing various other (and typically complex) methods for correcting bias that otherwise results when measurement error is ignored (Buonaccorsi, 2010; Culpepper & Aguinis, 2011; Ledgerwood & Shrout, 2011; Pieters, 2017).

Whether this tendency to neglect the effects of random measurement error reflects ignorance, laziness, or a lack of programming skill or familiarity with needed software to account for it during the analysis, it seems likely that the practice of ignoring measurement error in mediation analysis will continue without some further intervention. In this manuscript, we offer such an intervention by discussing and describing the implementation of a simple approach to managing the deleterious effects of measurement error in mediation analysis: errors-in-variables (EIV) regression. After a brief review of the mechanics of mediation analysis and the effects of random measurement error in estimation and inference, we describe EIV regression as an old, largely unused, but promising approach to overcoming the biasing of estimates of effects in a linear model that results from ignoring random measurement error. By way of three real-data examples and a small-scale simulation, we illustrate the effectiveness of EIV regression relative to ignoring measurement error or accounting for it using a more complicated single-indicator latent variable modeling approach using SEM. We end with a discussion of implementation of EIV regression in the freely-available PROCESS macro for SPSS, SAS, and R that now eliminates most any remaining excuses for not accounting for random measurement error when conducting a mediation analysis.

Mediation Analysis and Random Measurement Error

Although mediation analysis can take many forms depending on the nature of the variables and measurement systems used, the most common is the use of OLS regression analysis or SEM, with mediator M^O and outcome Y^O being continuous *observed* measurements (and hence the "O" superscript, as opposed to the theoretical *true scores* discussed later) of assumed-to-be continuous underlying constructs *M* and *Y*. In this scenario, the simplest mediation analysis is typically parameterized with a set of two equations, one for M^O and one for Y^O:

$$M_i^O = d_{M^O} + aX_i^O + e_{M_i^O} \quad (1)$$

$$Y_i^O = d_{Y^O} + c'X_i^O + bM_i^O + e_{Y_i^O} \quad (2)$$

where *a*, *b*, and c' are unstandardized regression weights, e_{M^O} are e_{Y^O} are errors in estimation of M^O and Y^O, respectively, d_{M^O} and d_{Y^O} are regression constants, and the *i* subscript denotes case, participant or observation $i = 1$ to *n* where *n* is the sample size. X^O can be a continuous observed measurement of construct *X*, just as M^O and Y^O, or it can be dichotomous numerical codes representing two groups, in the case of two-group experimental manipulation of X for example.[1] Modifications for multicategorical nominal or ordinal X^O variables are discussed elsewhere (e.g., Hayes & Preacher, 2014; Hayes, 2022, pp. 201-230) and beyond the scope of this paper, as is a discussion of models of dichotomous, count, or other forms of M^O or Y^O. Covariates can be added to the right-hand sides of equations 1 and 2 to deal with confounding of the effects of interest in a mediation analysis by shared causal influences on X^O, M^O, and/or Y^O. A visual representation of this model can be found in Figure 1, panel B.

A mediation analysis provides a quantification of three effects of X: the indirect effect, the direct effect, and the total effect of X. Of most relevance to mediation is the *indirect effect*, quantified as the product of a and b from equations 1 and 2. This, product, ab, estimates the difference in Y between two cases that differ by one unit on X resulting from the joint effect of X on M (estimated with a) which in turn affects Y (estimated with b). An indirect effect that is different from zero by some kind of inferential standard (a bootstrap confidence interval that does not contain zero being a popular inferential approach given the typical asymmetry of the sampling distribution of the product of two regression coefficients) provides evidence consistent with mediation of the effect on X on Y by M. Of course, mediation, as a causal process, cannot be established merely through a statistical analysis or examining the output of a statistical routine. Whether an effect can definitively be deemed causal requires strong theoretical argument and relevant design as much or even more than it does evidence than an effect, however quantified, is different from zero.

[1] Our words are carefully chosen here. X^O, M^O, and Y^O are a measures or manipulations of something but not necessarily what the researcher claims is being measured or manipulated. This paper is about the effects of and accounting for *random* measurement error, not about validity (i.e., whether the researcher is measuring and therefore studying what the researcher intends or claims). Whenever we used the term “measurement error,” we are talking about random measurement error.

The *direct effect* of X is estimated with c' in equation 2. It quantifies the difference in Y between two cases that differ by one unit on X but that are equal on M. It is everything about the effect of X on Y that is not carried through M. As such, it tells us nothing about mediation, as it captures the relationship between X and Y not attributable to mediation through M. The sum of the direct and indirect effect of X is the *total effect* of X, often denoted c. That is, $c = c' + ab$. The total effect estimates the average difference in Y attributable to a one-unit difference in X. There was a time when a mediation analysis would be undertaken only with affirmative evidence of an association between X and Y captured by c or a related statistic. But nowadays, it is understood that only ab (which is equivalent to $c - c'$ in single-mediator models estimated using regression analysis) is pertinent to mediation and this indirect effect can be different from zero even if c is not (see e.g., Hayes, 2009; Kenny & Judd, 2014; Shrout & Bolger, 2002; O'Rourke & MacKinnon, 2018). Thus, c and inference about it tells us nothing about mediation or whether X may be indirectly influencing Y through M.

This is typical practice in mediation analysis. But often, if not usually, researchers are more interested in the direct, indirect, and total effects of X estimated not from equations 1 and 2 but instead from

$$M_i^* = d_{M^*} + a^* X_i^* + e_{M_i^*} \tag{3}$$

$$Y_i^* = d_{Y^*} + c'^* X_i^* + b^* M_i^* + e_{Y_i^*} \tag{4}$$

as diagrammed in Figure 1 panel A, where X^*, M^* , and Y^* , are the are the underlying *true scores* of constructs X, M, and Y that the researcher is measuring. In this model, the direct, indirect, and total effects of X on Y are estimated as c'^*, a^*b^*, and $c^* = c'^* + a^*b^*$, respectively.

The distinction between true and observed scores is abstract but important and one with which most researchers are at least vaguely familiar. Recalling the example mediation analysis described at the beginning of this paper, Grisbook et al. (2024) used the Edinburgh Postnatal Depression Scale (EPDS) and the Psychiatric Diagnostic Screening Questionnaire (PDSQ) to measure postnatal depression and post traumatic stress. These self report measurement instruments generate observed scores that can be used in estimation of equations 1 and 2. But Grisbook et al. (2024) were interested in the effects of emergency C-section on child behavior through *depression* and *posttraumatic stress*, not through scores on the EPDS and the PDSQ. That is, they were interested in the effects estimated with equations 3 and 4, not equations 1 and 2.

The true scores are unobserved or "latent" and not directly quantified or available in the data, so equations 3 and 4 can't be directly estimated. Under classical test theory, a measurement theory motivating many measurement procedures in the behavioral sciences (see e.g., Nunnally, 1978), the observed scores are conceptualized as caused by the true scores (represented with the dashed arrows in Figure 1) but are not equivalent because the observed scores are a function of both the true scores and, typically, some random measurement error:

$$X_i^O = X_i^* + \varepsilon_{X_i}$$

$$M_i^O = M_i^* + \varepsilon_{M_i}$$

$$Y_i^O = Y_i^* + \varepsilon_{Y_i}$$

where ε_{X_i}, ε_{M_i}, and ε_{Y_i} are the random measurement errors for case *i*. These random errors in measurement can come from various sources that depend on the specifics of the measurement procedure or instruments being used, will vary between people, and can even vary in direction or magnitude over time and so can depend on when a person is measured. The point is that on any measurement occasion, the set of observed scores available in the data will not be the same as the true scores.

The amount of random measurement error that exists in a set of observed scores is the *reliability* of the observed scores, defined theoretically as the ratio of true score to observed score variance, the latter being the sum of the true score and random error variance under the assumption of classical test theory that errors are uncorrelated with true scores. For example, in the case of M^O:

$$\rho(M^O) = V(M^*) / V(M^O) = V(M^*) / [V(M^*) + V(\varepsilon_M)] \quad (5)$$

where V denotes variance and $\rho(M^O)$ is the reliability of M^O (reliability of X^O and Y^O are defined similarly). Reliability is the proportion of the variance in observed scores that is the result of variance in the true scores and so is between 0 and 1, with 0 reflecting observed measurements that are all random measurement error and 1 meaning an absence of random measurement error. Even though the true scores and therefore the variance of the true scores and random measurement error cannot be known, psychometric theory produces a variety of means of estimating the reliability of the observed scores, including Cronbach's α, McDonald's Ω, the correlation between observed measurements over time (test-retest reliability), and others.

When there is no random measurement error in the observed measurements, meaning reliability of all observed variables is 1, then $X_i^O = X_i^*$, $M_i^O = M_i^*$, $Y_i^O = Y_i^*$, and equations 1 and 3 and equations 2 and 4 are functionally equivalent. In other words, in the case of equivalence between the observed and true scores, $c' = c'^*$, $ab = a^*b^*$, and $c = c'^* + a^*b^*$ and so all is well when equations 1 and 2 are used to estimate the effects of interest in a mediation analysis.

But theory and research has shown that, typically, all is not well in a mediation analysis when observed X and/or M contain random measurement error. Such fallible measurement is quite common in research, but ignoring it when conducting a mediation analysis using equations 1 and 2 means that one or more of the effects, c', ab and c are likely to diverge from c'^*, a^*b^*, and c^*, with the extent of the divergence dependent on how much random measurement error exists in X^O and/or M^O, i.e., how far $\rho(X^O)$ and/or $\rho(M^O)$ deviate from 1. That is, the result of using such fallible measurements and estimating equations 1 and 2 will be inaccurate estimates of one or more of the effects of interest, c'^*, a^*b^*, and/or c^*, and inferential tests for those effects from computations based on equations 1 and 2 that are likely to be invalid. For a discussion and evidence, see Cheung and Lau (2008); Cole and Preacher (2014); Fritz, Kenny, and MacKinnon (2016); Gonzales and MacKinnon (2021); Valeri, Lin, and VanderWeele (2014); VanderWeele, Valeri, and Ogburn (2012).

Note that ignoring measurement error is not a death sentence for accuracy of estimation and test validity. If Y^O contains random measurement error but X^O and M^O are perfectly reliable, the *unstandardized* effects (the effects we are focusing on in this paper) estimated by equations 1 and 2 will be estimating the same thing that c'^*, a^*b^*, and c^* in equations 3 and 4 do.[2] However, sampling variance of these effects will be larger, meaning power to detect the direct, indirect, and total effects will be reduced and confidence intervals will be wider than they would be if Y^O were free of random measurement error.

Approaches to Managing the Effects of Measurement Error

Though we have no data to support this claim, we believe anyone who is familiar with the empirical research in their area would agree that the typical practice for mitigating the effects of random measurement error in a mediation analysis is to do nothing. It is apparent from the use of popular tools that simplify a mediation analysis such as the PROCESS macro for SPSS, SAS, and R (Hayes, 2022) and its precursors (Preacher & Hayes, 2004, 2008), the mediation package in R (Tingley, Yamamoto, Hiros, Imai, & Keele, 2014), PROC CAUSALMED in SAS, and other computational aids that historically or currently have no means of incorporating measurement error into the estimation that most published mediation analyses contain some and perhaps substantial bias as a result, along with the corresponding effects of such bias on the accuracy and validity of inference (c.f., Cole & Preacher, 2014).

That said, this strategy, if it can be called that, is not always quite as problematic as it might seem. As mentioned earlier, random measurement error in Y^O does not by itself bias the estimation of effects when using equations 1 and 2, and when X is experimentally manipulated or codes groups on well-defined attributes, the investigators usually at some point knows with certainty which

---

[2] This is true for unstandardized effects. Accuracy in estimation of standardized effects will be influenced by random measurement error in Y^O as well. A discussion of the effects of measurement error on standardized paths and effects in a mediation analysis is beyond the scope of this paper.

condition a participant was assigned to or which category a person is a member of, meaning the reliability of X^O can be assumed to be one.[3] That leaves measurement error in M^O (as well as any covariates) as the remaining source of measurement error-induced bias in estimation. If the investigator is careful to minimize random measurement error in M^O (and covariates), the bias is likely to be smaller, though not zero, than it otherwise would be when X^O is not the same as X^*.

Latent Variable SEM

If an attempt is made to acknowledge random measurement error when conducting a mediation analysis, most often the investigator will choose SEM. Using an SEM program for estimation does not in of itself necessarily do anything to address the problem if the model being estimated is just a path analysis linking observed variables that contain random measurement error together in a structural model. Rather, the model must include some kind of measurement model component in addition to the structural component that keeps the random measurement error out of the mathematics that generates the structural path coefficients between latent variables (the SEM-equivalent of the "true scores."). This can be done using either a multiple indicator latent variable (MILV) approach or a single-indicator latent variable (SILV) approach.

The MILV approach can be used when one or more of the variables *X*, *M*, and/or *Y* in a mediation model is measured with two or more indicators of the underlying latent variable. Indicators might be, for example, the individual questions on a self-report measure that a person is asked to respond to during the measurement procedure. Measurements of these indicators are observed and therefore available in the data and specified as causally influenced by the latent variable, which is unobserved and so not in the data. The quantification of the effect of the latent variable on an indicator is the indicator's *factor loading*. The model of each indicator also includes an error in estimation whose variance is estimated and that captures both random measurement

[3] An exception would be when an investigator creates a categorical variable by artificially categorizing a continuous variable. In this case, the reliability of the categorical variable will be less than one (MacCallum, Zhang, Preacher, and Rucker, 2002)

error and error in the estimation of the indicator variable from the latent variable (as rarely would an indicator be perfectly predictable from the latent variable causally influencing it). These linkages between latent variables and their indicators constitute the measurement component of the model. The structural component of the model is the paths, assumed to be causal in a mediation model, that connect the latent variables together in a causal system, just as in an observed variable model. This procedure keeps the random measurement error on the measurement side of the model and will produce more accurate estimates of the effects of the latent variables in the mediation analysis. Examples of the MILV approach in action in a mediation analysis can be found in Fye, Kim, and Rainey (2022) and Wang, Sang, Li, and Zhao (2016). Mediation analysis using the MILV approach is discussed in more detail by Cheung and Lau (2008), Falk and Biesanz (2015), Lau and Cheung (2012), and MacKinnon (2008). Bollen (1989) and Kline (2023) discuss the theory and practice of SEM and latent variable modeling.

When using the MILV approach, no information is required about the reliability of an observed variable. By contrast, the SILV approach involves the estimation of the effects in a structural model just as in the MILV approach, but the measurement component of the model is simpler, with the observed variables X, M, and/or Y being the sole indicators of their corresponding latent variables, as in Figure 2. As the measurement model for each variable won't be identified when there is only one indicator, constraints must be imposed that both identify the measurement side of the model and specify the reliability of the single-item indicators of the latent variables, i.e., the reliability of the observed variables. This is accomplished by fixing the effect of the latent variable on the observed variable to 1 and the variance of the random error in the observed variable to the variance of the observed variable multiplied by 1 minus an estimate of the reliability of the observed variable. Thus, unlike the MILV approach, the SILV approach requires an external estimate of the reliability of the observed variables. The important point is that whether using the SILV or MILV approaches, the variance-covariance matrix for the latent variables used in the estimation of the structural paths in the mediation model is less contaminated by measurement error than it otherwise would have been if the reliability of the observed variables were assumed to be one and the model estimated as an observed variable SEM or using separate OLS regression equations.

Errors-in-Variables Regression

An alternative and arguably simpler approach to accounting for the effects of measurement error, EIV regression is often attributed to Fuller (e.g., Fuller, 1987) and has origins in the economics literature. This approach is both elegant and not terribly difficult to implement (see Appendix A for the EIV computations we use throughout this manuscript). The premise is that regression coefficients are derived from the variance-covariance matrix of the variables in the model, but the variance-covariance matrix of observed scores is contaminated by measurement error. Under the assumptions of classical test theory that the random measurement errors and true scores are uncorrelated and the random errors are uncorrelated with each other, the covariance between observed scores is equal to the covariance between the true scores. Thus, the covariances are fine as is even with random measurement error in the observed scores. But the variances of the observed scores contain a component attributable to random measurement error.

EIV regression involves a modification of the variances of variables on the right-hand sides of the equations defining a model that removes the variance of the observed scores attributable to random measurement error. Under the laws of classical test theory, the variance of the random measurement errors for the mediator is, from equation (5),

$$V(\varepsilon_M) = [1 - \rho(M^O)]\, V(M^O) \tag{6a}$$

Using a similar logic for *X*,

$$V(\varepsilon_X) = [1 - \rho(X^O)]\, V(X^O) \tag{6b}$$

and so the variance of the mediator and *X* true scores are, respectively,

$$V(M^*) = V(M^O) - [1 - \rho(M^O)]\, V(M^O) \tag{7a}$$

$$V(X^*) = V(X^O) - [1 - \rho(X^O)]\, V(X^O) \tag{7b}$$

With information about the reliability and variance of the observed scores for *X*, *M*, and covariates if used, the variance of the true scores can be calculated. Reliability estimates are not difficult to generate using methods discussed in the psychometrics literature, taught in many methodology classes researchers take in graduate school and elsewhere, and are programmed into most good statistics packages, and of course the variance of the observed scores is known. Since the input to the regression routine is the variance-covariance matrix of the variables in the model, EIV regression works with a version of the variance-covariance matrix whose diagonal elements for the variances of the variables on the right-hand side of the equation are produced by subtracting from the observed variances the part of the variance in the observed scores due to random measurement error, as in equations 7a and 7b (and likewise for covariates if used).

So long as the estimate of reliability of the observed scores is reasonably accurate, in effect this modification produces a data matrix that corresponds more closely to the data one would have if the true scores rather than the observed scores had been quantified. The new data matrix is then used with standard OLS regression algebra to produce estimates of the effects of variables in the model that are likely to be closer to their "true" or correct values had the true scores been observed instead of the error-contaminated observed scores. Because of the modification to the data, standard errors for the resulting regression coefficients must be adjusted to account for this modification (see Appendix A). Note that in EIV regression, random measurement error in the variable on the left-hand side of a model equation is not removed from the data, as random measurement error in that variable when serving as outcome does not bias the estimation of regression coefficients.

Methodologists have examined the utility of EIV regression when analyzing social science data. For example, Culpepper and Aguinis (2011) found that EIV regression produced more accurate estimates of mean differences while maintaining Type I error control compared to OLS regression and a few alternative approaches to random measurement error correction in the two-group analysis of covariance when the covariate contains random measurement error. And Culpepper (2012) provides evidence of the superior performance of EIV regression compared to OLS regression and the SILV approach for testing interaction between two variables when one is dichotomous. Counsell and Cribbie (2017) also found good performance of EIV regression relative to change scores or analysis of covariance in two-group comparisons of change over time, but they were less optimistic about is use in some circumstances such as when the sample size was small or the assumed reliability of the variable over time was different than the actual test-retest reliability.

Some exceptions aside, EIV appears to have promise as a straightforward solution to the problems in estimation of effects in linear models that result when measurement error is ignored. But the impact of this work on the practice of data analysis, mediation analysis in particular, has thus far been limited as evidenced by the lack of use of EIV regression in the behavioral science literature. Two explanations for this seem probable. First, EIV regression likely remains largely unknown to most researchers, as is not implemented in most data analysis software behavioral scientists prefer to use (the exception being Stata) nor is it discussed much if at all in popular books often used in regression and linear modeling classes. Second, the work to date on EIV regression has not been undertaken with mediation analysis in mind and so perhaps has not attracted the attention of those who otherwise have taken great interest in elucidating mechanisms that underly causal effects through mediation analysis.

In the rest of this manuscript, we explore and discuss the potential contribution of EIV regression as an approach to managing the effects of random measurement error in mediation analysis. We first do a set of example mediation analyses, analyzing real data sets using OLS regression analyses of observed variables as well as using EIV regression. We do the same analyses using the SILV approach to compare the results it yields relative to OLS and EIV regression. We then describe a simulation that addresses better than our examples whether EIV reduces or eliminates the bias observed when random measurement error is not accounted for using OLS regression. We end with a discussion of an implementation of EIV regression in the PROCESS macro for SPSS, SAS, and R, acknowledging some potential limitations and caveats of using EIV regression.

Example Mediation Analyses Using Real Data

In this section, we provide three example mediation analyses using real data sets from published studies, each of which contained a mediation analysis. We analyzed the data the investigators made publicly available using three approaches, thereby allowing a comparison of point estimates and inferences they yield: OLS regression which does not account for random measurement error, EIV regression accounting for random measurement error in variables on the right side of equations, and the SILV modeling approach. Code used to conduct each analysis can be found in Appendix B.

Example 1: Compassion Fatigue and Compassion Mindset

The first example is based on data taken from Gainsburg and Cunningham (2023). The data include responses from 308 adults in the United States who completed a task designed to elicit compassion fatigue by showing them photos of people experiencing distressing situations. Participants were asked about their beliefs about compassion as a limited resource, expected compassion fatigue from the task, and resulting compassion fatigue. The mediation model specifies that beliefs about whether compassion is a limited resource (compassion mindset: X) as the cause of compassion fatigue elicited from the task (experienced compassion fatigue: Y), operating indirectly through their anticipated compassion fatigue from the task (expected compassion fatigue: M). Participants' expectations about how fatiguing the task will be were theoretically caused by their beliefs about compassion as a limited resource, which in turn would increase their experience of compassion fatigue. Measures of all three variables were constructed as unweighted averages of responses to multiple indicators of each construct in measured using a Likert response format. The reliabilities of each construct provided by the investigators and using Cronbach's α were 0.73 for compassion mindset (four indicators), 0.86 for expected compassion fatigue (eight indicators), and 0.84 for experienced compassion fatigue (eight indicators).

A mediation analysis was conducted first using ordinary least squares regression models of expected compassion fatigue (*M*: Equation 1) and experienced compassion fatigue (*Y*: Equation 2), which ignores random measurement error in all three variables. The analysis was conducted using PROCESS (Hayes, 2022) with inference for the indirect effect conducted using a percentile bootstrap confidence interval based on 5,000 bootstrap samples. Other tools capable of conducting a mediation analysis could be used, such as structural equation modelling of observed variables using maximum likelihood estimation. SEM models would produce similar estimates (Hayes, Montoya, & Rockwood, 2017), with slight differences in standard errors resulting from the use of maximum likelihood.

We next used EIV regression as implemented in the PROCESS macro as of Version 5 that is described in more detail later. The code to estimate the model can be found in Appendix B and the mathematical details of the implementation can be found in Appendix A. EIV regression accounts for random measurement error on the right-hand sides of each equation but not the left-hand side. In the model predicting expected compassion fatigue (*M*), random measurement error was accounted for compassion mindset (*X*) during model estimation through modification of the data matrix to account for random measurement error in *X* but not in *M*. In the model predicting experienced compassion fatigue (*Y*), the model was estimated using a modified data matrix accounting for random measurement error in compassion mindset (*X*) and expected compassion fatigue (*M*). Inference for the indirect effect was conducted using percentile bootstrap confidence intervals based on 5,000 bootstrap samples.

The model was then estimated using the SILV modeling approach as diagrammed in Figure 2 and using the lavaan package version 0.6-18 in R (code is provided in Appendix B, with accompanying code for the same analysis in Mplus and Stata available in supplementary materials available upon request). The unweighted average of indicators for each variable served as the sole indicator of their respective latent variables, with the factor loading constrained to one and the error in the estimation of the indicator variables constrained to the observed variance multiplied by one minus the reliability estimate. The mediation model was then estimated replacing the observed variables with corresponding latent variables using maximum likelihood estimation of model parameters, and inference about the indirect effect estimated using a percentile bootstrap confidence interval based on 5,000 bootstrap samples.

The results of these three analyses are shown in Table 1. Notice first that the OLS estimates of each path and the indirect effect differ from those generated by the EIV and SILV approaches, with the OLS estimates attenuated (i.e., closer to zero) relative to those generated with the EIV and SILV approaches. This attenuation of effects is consistent with previous literature on the effects of measurement error on the estimates of effects in linear models. But as discussed in Cole and Preacher (2014), the consequences of unaccounted-for unreliability can be either the under or overestimation of effects depending on the complexity of the model, which variables are measured imperfectly, and how unreliable those measurements are (as our second example will illustrate).

Second, notice in Table 1 that the point estimates of the effects using EIV regression and the SILV approach are largely the same. In this example, it appears to make little difference which of the two methods is used. The point estimates are not affected by the choice. We would expect that the standard errors would be smaller using the SILV approach, as it accounts for random measurement error in variables on the left sides of the equations. However, we don't see a pattern consistent with this expectation. Perhaps the sample size is large enough in this example to eliminate any of the precision of estimation advantages that would come with accounting for measurement error in the outcome.

Finally, observe that substantively, and thinking only dichotomously in terms of whether an effect can be said to be zero or not, the results are very similar between the three approaches. All three yield 95% confidence intervals for the indirect effect that are positive and exclude zero, and both result in total and direct effects that are not statistically different from zero by a null hypothesis test or confidence interval.

Example 2: Nature and Self-Actualization

The second example is based on data from Yang et al. (2024, Study 4), who investigated the effect of exposure to nature on authenticity, defined as a sense of humanistic self-actualization with behaviors congruent with the self. One hundred seventy one participants were randomly assigned to view either pictures of nature ($X = 1$) or pictures of urban environments ($X = 0$), then answered questions about their resulting mood (positive affect: M) and sense of authenticity (authenticity = Y). It was hypothesized that participants in the nature condition relative to those in the urban condition would have a greater sense of authenticity (12 indicators averaged to produce a composite measure, Cronbach's $\alpha = 0.82$) indirectly through an enhancement of positive mood (18 indicators averaged, Cronbach's $\alpha = 0.91$), which would intern prompt greater authenticity.

The mediation analyses were conducted just as described in the first example (accompanying code is shown in Appendix B), with results from the three analytical approaches shown in Table 2. Models predicting M are largely identical across approaches, since the experimental manipulation contains no random measurement error and measurement error in the variable on the left-hand side of the equation (M) does not bias the estimate of X's effect. The differences between approaches become apparent when looking at the estimates in the equation for Y. Like the previous example, the OLS estimate of the effect of the mediator on Y is attenuated relative to the EIV and SILV estimates, but the direct effect of X is closer to zero in the EIV and SILV models compared to the OLS model. Given that the effect of X on M and the total effect of X on Y is same across approaches (a requirement of an X without measurement error), the attenuation of the effect of M on Y in OLS means that the direct effect of X would have to increase to offset the resulting decrease in the indirect effect relative to the EIV and SILV approaches. While all three approaches produce statistically significant indirect effects and non-significant direct effects, the magnitude of the indirect effect is larger and direct effect is subsequently smaller in EIV and SILV models. And like in the first example, the EIV and SILV approaches produce largely identical estimates of the effects and only trivially different standard errors, but the estimates are different from those produced by OLS regression.

Example 3: Photo-Editing and Self-Perceived Attractiveness

The final example is based on data taken from Ozimek et al. (2023). The data includes responses from 316 adults who actively use social media about their photo editing behavior and self-perceptions. Photo editing (X) was measured using the unweighted average of five indicators measuring the participant's tendency to edit and use filters on pictures of themselves for social media, with higher scores representing more photo editing behavior (Cronbach's $\alpha = 0.75$). Participants were also asked about their self-objectification (M), a 14-item measure quantifying the degree to which participants view themselves and their own self-worth through their attractiveness as perceived through the eyes of others (Cronbach's $\alpha = 0.89$). Participants were also asked to rate their self-perceived attractiveness (Y) using six indicators assessing their general satisfaction with their appearance (Cronbach's $\alpha = 0.94$). They proposed that higher photo editing behaviors contribute to lower self-perceived attractiveness indirectly through increases in self-objectifying which would in turn lower self-perceived attractiveness.

The mediation analyses were conducted as in the previous two examples, and the results of are shown in Table 3. In line with the previous two examples, EIV and SILV approaches produce similar estimates of regression coefficients, total, direct and indirect effects. And as in the first example also based on an imperfectly measured X and M, OLS estimates are attenuated relative to the other two approaches, but statistical inference yields the same conclusions about which effects are different from zero and which are not.

The Effectiveness of EIV Regression: Simulation Evidence

Our example analyses presented in the prior section show that EIV regression produces estimates of the effects in a mediation model that correspond closely to the estimates obtained from the SEM-based SILV approach, both of which diverge from estimates generated by OLS regression. But two people can agree the earth is flat while both being wrong. Who is to say that the EIV and SILV approaches generate more accurate estimates and inferences? Without some objective truth against which these results can be evaluated, the fact that EIV results correspond to the SILV model results doesn't mean the EIV results are more trustworthy. Perhaps they are both wrong in the same manner and the OLS results better estimates of reality. To answer this question, we conducted a small-scale simulation in which we defined truth to determine whether the EIV approach that accounts for random measurement generates more accurate estimates of the effects in a mediation model than does OLS regression that ignores that random measurement error.

We assume that researchers who ignore random measurement error by conducting a mediation analysis using OLS regression are comfortable treating their point estimates of effects and the fit of their models (using R^2) as estimates of corresponding parameters in the population from which they have sampled or the true data generating mechanism. Although we have reason to doubt that the OLS estimates are good ones, we designed the simulation giving the benefit of the doubt to the OLS regression results we report in Tables 1-3. We treated the OLS point estimates of all the paths in those models and the resulting direct, indirect, and total effects as population values or "parameters" of the population or data generating mechanism. Furthermore, we treated the squared multiple correlations as population values of variance explained in M and Y by each equation. We used these parameters in our simulations, creating data sets of sizes $n = 50$, 100, 200, 300, 500, and 1000 that represent random samples from population mediation models defined by the OLS regression estimates in Tables 1-3. For the simulation corresponding to examples 1 and 3 with continuous X, the data generation started with a set of random standard normal deviates for X, whereas for example 2, X was a dichotomous variable coded 0 and 1 for the two groups with the sample split equally between the two groups. The errors in estimation in the models of M and Y were random normal deviates with variance set to the value required to produce the corresponding squared multiple correlation (within expected sampling error) for each model equation.

This first stage of data generation just described produced a trivariate sample from the population with X, M, and Y containing no random measurement error, i.e., representing true scores on the latent variables, with relationships between them defined by the population mediation model. Next, we added random normal measurement error onto X (except in the simulation with a dichotomous X), M, and Y such that the resulting observed scores had reliabilities equal to (within expected random sampling error) the reliabilities reported in our example analyses. Using the resulting trivariate sample, now adulterated by random measurement error, we estimated the direct, indirect, and total effects of X using OLS regression and EIV regression, specifying the same reliabilities in the EIV regression routine that were used to generate the observed data. This procedure was repeated for a total of 5,000 times for each sample size, recording each of the estimated effects in each repetition as well as whether a 95% confidence interval for the effect included the known population value. As the sampling distribution of an indirect effect is not normal in form, confidence intervals for the indirect effect were generated using the percentile bootstrap method based on 5,000 bootstrap samples, whereas confidence intervals for the direct and total effects were generated in the usual way, assuming the sampling distribution is roughly normal in form, as the point estimate plus or minus approximately 2 standard errors (the appropriate critical value from the t distribution was used rather than 2, with degrees of freedom equal to the residual degrees of freedom for the regression equation). We repeated this entire simulation for a total of 18 times, each time using one of the sets of population parameters defined by the models reported in Tables 1, 2, and 3 and for 6 sample sizes. All simulations were programmed in R and using the OLS and EIV regression routines built into the PROCESS macro described later as the computational engine for estimation of the model coefficients and calculating inferential statistics. The PROCESS implementation of EIV regression is documented in Appendix A.

The results are found in Tables 4 and 5 for each population and sample size combination. Table 4 provides mean estimate of each effect over the 5,000 replications, as well as the mean bias percentage, defined as Bias% = (Mean Estimate – Parameter) / Parameter. The value of the parameter of each effect is found in each of the subheadings. A negative value for Bias% reflects attenuation of the effect toward zero, whereas a positive value reflects an overestimate (i.e., bias away from zero). Table 5 provides 95% confidence interval coverage, meaning the percentage of times over the 5,000 replications that the confidence interval for that effect included the parameter being estimated. Good performance is reflected in a mean estimate of the effect close to the corresponding parameter (i.e., Bias% near zero) and confidence interval coverage near 95%.

Given that we have simulated only three populations varying unsystematically in the sizes of the effects being estimated and the reliabilities of the observed variables, we are cautious to not overanalyze and overinterpret these results. But there are a few patterns that jump right off the page and are noteworthy. First, notice in Table 4 that OLS generally gets the effects wrong, and sometimes substantially so, as reflected by the difference between the population effects and the mean estimates of those effects over the 5,000 replications. And observe that increasing the sample size has little to no effect on the bias in OLS estimates. You can't make the problem produced by unaccounted-for random measurement error go away or even diminish by just collecting more data. But EIV regression generally gets it right regardless of sample size, with very little discrepancy between mean estimates and the population values of the effects.

Second, turning to Table 5, notice that confidence interval coverage using EIV is generally right on the money, with about 95% of 95% confidence intervals containing the parameter, plus or minus a few percent here and there. Not so for OLS regression-based confidence intervals, but with a caveat discussed later. Notice that in a few of our example populations, OLS confidence interval coverage for the indirect effect is below 95% even in small samples, and the larger the sample, the worse things get. This is the result of the bias observed in Table 4. As the sample size increases, the confidence interval becomes increasingly narrow, converging around the *wrong* estimate and increasingly excluding the parameter. In smaller samples, the bias is likely offset by the larger confidence interval width such that the interval is more likely to capture the parameter even though the interval is centered around a biased estimate.

Earlier we mentioned that random measurement error is not a death sentence for accurate estimation and inference. The second simulation based on the nature and self-actualization study makes this point. In this simulation, *X* is dichotomous and contains no random measurement error and *M*, though not free of measurement error, is measured with fairly high reliability (0.91). Measurement error in *M* but not *X*, regardless of the extent of measurement error in *Y*, will tend to bias the estimation of the effect of *M* on *Y* and therefore the indirect effect of *X* while also influencing to some extent the accuracy in the estimation of the direct effect of *X*. This is what we see in the second simulation. But it would not bias the estimation of the total effect, as seen in Table 2. For this to happen, the biases in the estimation of the direct and indirect effects would have to be similar in magnitude but opposite in direction given that the total effect is the sum of the direct and indirect effects. This is also what we see in Table 4. But given the high reliability of *M* in this example, the bias in estimation of the direct and indirect effects is small, and confidence coverage is still respectable even in moderate to large samples.

In summary, the evidence from this limited simulation is consistent with the conclusion that OLS estimation is the flat-earther here and that EIV regression (and, by logical extension given the example analyses presented in the prior section, the SILV approach) correctly sees the world as round. This conclusion agrees with research and analytical derivations in other modeling contexts that random measurement error in variables on the right-hand side of linear models can wreak havoc on accurate estimation of the effects the researcher is trying to quantify and test.

An Easy-to-Use Implementation of Errors-in-Variables Regression

Fortunately, EIV regression is not some obscure computational technique that one must have a degree in statistics or computer science to employ. It is already available in Stata (StataCorp, 2023) as well as in R (Culpepper and Aguinis, 2011; with code later updated at www.hermanaguinis.com/eiv.html), including in the "eivtools" package available through CRAN. These implementations differ and will produce slightly different results and have different output options even though they are based on roughly the same statistical theory. But these implementations were not created with mediation analysis in mind and so require the user to estimate all the equations for a mediation model with separate commands, and obtaining inferential tests for indirect effects requires additional programming that likely goes beyond the skills of many researchers. To facilitate adoption of EIV regression in mediation analysis (and linear modeling more generally) by easing the programming and computational burden, we instead recommend the easy-to-use PROCESS macro for SPSS, SAS, and R (Hayes, 2022) used in our examples and simulations. PROCESS is freely-available at *www.processmacro.org* and already enjoys wide use in mediation analysis throughout the behavioral sciences. EIV regression is implemented as of version 5 and can be used for any mediation model PROCESS can estimate, including simple (single mediator), multiple (parallel or serial), blended (combining parallel and serial) and custom mediation models programmed as described in Appendix B of Hayes (2022).[4] Covariates with corresponding reliabilities can also be included in any mediation model. Only a single line of code is required rather than lengthier syntax required when using an SEM framework. SPSS users have the option of setting up the model with a user-friendly graphical user interface if desired. PROCESS takes care of all the computational work, including inference for indirect effects using bootstrapping. PROCESS implements the computations described in Appendix A. How to set up

[4] The EIV regression routine in PROCESS is not yet available for moderation models or models that combine moderation and mediation (conditional process models).

PROCESS and execute a PROCESS command is documented extensively in Hayes (2022), though the documentation there does not include any discussion of the EIV routine as it was implemented after the printing of the book. Below we provide a brief discussion of the EIV regression options available in PROCESS.

For a mediation analysis, PROCESS will expect the user to specify a single (and only one) independent variable *X* following **x=**, a single outcome or dependent variable *Y* following **y=** , and at least 1 mediator *M* following **m=**. Covariates can be included if desired using the optional **cov=** option. Reliabilities for *X* and *M* are entered using the **relx=** and **relm**= options. A model number is typically also required. Using **model=4** in the PROCESS command specifies a simple (single mediator) or parallel multiple mediator model, whereas **model=6** specifies a serial multiple mediator model. Models 80, 81, 82 are models that blend parallel and serial mediation (see the documentation).

In Appendix B, we provide the PROCESS command for the SPSS, SAS, and R versions of the example analyses presented earlier using EIV regression. As can be seen there, the PROCESS command for the compassionate fatigue study specifies the variables in the data named “compass” as *X*, “efatigue” as *M*, and “fatigue” as *Y*, with the estimated reliabilities of compass and efatigue being 0.73 and 0.86, respectively. The model=4 option tells PROCESS to set this up as a mediation model. The resulting PROCESS output can be found in Appendix C. Notice that PROCESS provides model summary information, the regression coefficients, standard errors, *t*- and *p*- values, confidence intervals, and estimates of the direct, indirect, and total effects of *X* with corresponding inferential information for those effects.

Consider a modification to this example analysis that includes an additional mediator in the model named “rational” in the data and measured with reliability 0.82. In addition, suppose we wanted to include three covariates: “age” in years, whether or not a person self-identified as “male” (coded 1 in the data, 0 otherwise), and self-esteem (“selfest”) measured with reliability 0.95.

Assuming that age and whether a person identifies as male were measured without random measurement error (i.e., with reliabilities equal to 1), the PROCESS command below estimates a mediation model with “affect” and “rationale” operating as parallel mediators:

SPSS

```
process y=fatigue/x=compass/m=efatigue rational/cov=age male
  selfest/model=4/relx=0.73/relm=0.86,0.82/relcov=1,1,0.95.
```

SAS

```
process (data=compfat,y=fatigue,x=compass,m=efatigue rational,cov=age
  male selfest,model=4,relx=0.72,relm=0.86 0.82,relcov=1 1 0.95)
```

R

```
process(data=compfat,y="fatigue",x="compass",m=c("efatigue",
  "rational"),cov=c("age","male","selfest"),model=4,relx=0.72,
  relm=c(0.86,0.82),relcov=c(1,1,0.95))
```

When more than one mediator or covariate is listed following **m=** or **cov=**, reliabilities, if any, must be provided for all variables in the same order the variables are listed in the PROCESS command. Unknown reliabilities could be set to 1 or the user could provide a reasonable guess for the unknown reliability if assuming 1 is not defensible. If the reliability of the measured variables is not provided by using the **relx**, **relm**, or **relcov** options (if covariates are included in the model), reliability for those variables is treated as 1 by PROCESS. In other words, reliability of 1 is the default. When neither **relx**, **relm**, nor **relcov** options are used in the command, PROCESS does an OLS regression analysis rather than EIV regression.

In a mediation analysis, PROCESS defaults to the production of a 95% bootstrap confidence interval for the indirect effect(s) using the percentile method based on 5,000 bootstrap samples. The number of bootstrap samples can be changed with the **boot** option (e.g., **boot=10000** for 10,000 bootstrap samples) and the confidence level changed with the **conf** option (e.g., **conf=90** for 90% confidence intervals). Bootstrap inference for each of the regression coefficients in the model is also available rather than for just the indirect effect(s) by using the **modelbt** option. Output can also be saved for use later if desired using the **save** option, which we relied on to conduct the simulation reported earlier. See the PROCESS documentation in Hayes (2022) for a discussion of these options.

EIV regression relies on a modification of the data based on the reliabilities of the observed variables provided by the user. In some circumstances, the resulting modified variance-covariance matrix may not be possible to observe in nature. This can occur in small samples and/or when one or more of the entered reliabilities is too small. In that case, PROCESS will produce an error saying the model could not be estimated as a result of one or more small reliabilities entered. That can also occur during the bootstrapping phase. When it does, PROCESS will replace the offending bootstrap sample with another. A note at the bottom of the output will tell the user how many times a bootstrap sample had to be replaced. Although there is no guidance available for how many such replacements is acceptable without affecting the validity of bootstrap inference, common sense would suggest fewer replacements would be better. If the user is uncomfortable with the number of replacements required to complete the bootstrapping, a Monte Carlo confidence interval could be used as an alternative for inference about the indirect effect, as it does not require resampling from the data (Preacher & Selig, 2012). PROCESS can conduct a Monte Carlo confidence interval for indirect effects using the **mc** option, as described in the documentation.

Note that a mediator is not required to use PROCESS's EIV regression option. It will also conduct an ordinary EIV regression analysis. For example, using the same variables and reliabilities from the prior hypothetical example, the command below would estimate an EIV regression model of fatigue from compass, efatigue, rational, age, male, and selfest:

<u>SPSS</u>

```
process y=fatigue/x=compass efatigue rational age male selfest/
  relx=0.72,0.86,0.82,1,1,0.95.
```

<u>SAS</u>

```
process (data=compfat,y=fatigue,x=compass efatigue rational age male
  selfest,relx=0.72 0.86 0.82 1 1 0.95)
```

<u>R</u>

```
process(data=compfat,y="fatigue",x=c("compass","efatigue","rational",
  "age","male","selfest"),relx=c(0.72,0.86,0.82,1,1,0.95))
```

Because analysis of covariance (ANCOVA) is just a special case of multiple regression, it follows that the EIV routine in PROCESS can be used to conduct analysis of covariance when the covariate or covariates contain random measurement error (see Culpepper and Aguinis, 2011, for a discussion of EIV regression in ANCOVA). Using the **mcx** option described in the documentation, the user can specify that there are more than two groups begin compared in the ANCOVA.

Discussion

In this manuscript, we have made the case for EIV regression as a viable approach to mediation analysis when the variables in the model contain random measurement error. Our example analyses and simulations show that EIV regression produces results comparable to the single indicator latent variable approach using SEM while reducing or eliminating the bias in the estimation of effects that occurs when using OLS regression and ignoring random measurement error. Furthermore, at least in our examples, EIV-based confidence intervals for effects preserved their meaning as coverage was generally consistent with confidence. But OLS confidence intervals lose their meaning with increasing sample size, as coverage of the true effect decreases (gets worse) with more data as a result of the estimation bias. We recommend that researchers properly address the random measurement error in their mediation analyses, with EIV-regression being a rather painless and easy-to-use approach, even for more complicated models than we have focused on here, and certainly better than the standard practice of ignoring it altogether.

Although we encourage researchers to use EIV regression when conducting a mediation analysis, it could also be used as supplementary method rather than the main analysis reported. For example, PROCESS makes it easy to conduct a sensitivity analysis for results generated when random measurement error is ignored or different than assumed. Perhaps the investigator prefers for one reason or another to use OLS regression but wants to know how vulnerable the conclusions are to unaccounted-for measurement error. The investigator could repeat the mediation analysis that ignored measurement error but using known estimates of reliability in an EIV routine such as programmed in PROCESS. Consistency in the results, at least substantively, can then be used as a defense against criticism that the results reported when unreliability is ignored are just artifacts of unaccounted-for measurement error. Another possibility is to follow up an EIV-based mediation analysis with alternative reliability estimates that are smaller than the reliabilities used. This would assess the vulnerability of the original analysis to random measurement error that is larger than assumed. Or if there is no estimate of reliability available for one or more variables in the model, the investigator could try different values of reliability to see how much the results are affected by different assumptions about the unknown reliability.

In the realm of pedagogy, we hope our findings about the viability of EIV regression-based mediation analysis and its implementation in PROCESS will encourage statistics and methodology instructors to address the shortcomings of unaccounted-for measurement error in data analysis and how it can be quite easily addressed in some circumstances. Contrary to the difficulties many students experience learning to use an SEM program, we have found that it takes only 10-15 minutes of classroom time to get students familiar with PROCESS syntax, and its implementation in the graphical user interface in SPSS makes it still easier to employ, even in undergraduate classrooms. The speed at which PROCESS generates output is an additional advantage of its use in the classroom relative to lavaan and Stata.

But an analysis is only as good as the data it is given. By the standard of reliability, the less well a variable in a model is measured, the less well that variable will capture the effect of what the researcher claims to be studying. The same can be said for any strategy that attempts to account for the effects of random measurement error. EIV regression, like the SILV approach, requires information about the reliability of the observed variables in the model. In the example analyses we reported here, we treated the reliability estimates reported by the original investigators as truth. And in the simulations, we defined that truth. But in practice, investigators don't know the reliability of their measurements. They can only estimate their reliability, and there are different approaches to estimation that can give different estimates. To satisfy its mission of accounting for the effects of random measurement error, EIV regression must be given good data to make the adjustment, and that means providing the algorithm with reasonably accurate reliability information. Had we used different reliability estimates in our examples or used incorrect reliability information in our simulation, EIV regression would have gotten things wrong too. As Culpepper (2012) and Counsell and Cribbie (2017) note in a different analytical context, the debiasing effects of EIV-regression are dependent on providing the routine with accurate estimates of reliability of the observed variables. So investigators should take care in generating those estimates.

Lockwood and McCaffrey (2020) report that EIV standard errors may be too small in some circumstances, and Counsel and Cribbie (2017) reported elevated Type I error rates in smaller samples for inferences about group differences in a two-group pretest-posttest design. Inaccurate standard errors would not affect the validity of inferences for the indirect effect using a percentile bootstrap confidence interval, as no standard error is used in the derivation of a bootstrap confidence interval calculated using the percentile method. But it could affect the validity of inferences in small samples about individual paths, including the total and direct effects. To check on this, we calculated the average EIV standard error estimated using expression A4 described in Appendix A (the one Lockwood and McCaffrey studied and similar to though not identical to the one that Counsel and Cribbie, 2017, reported using), the standard error estimator we used in our simulation (expressions A2 and A3 in Appendix A), as well as the average standard deviation of the bootstrap EIV estimates of each path (a^*, b^*, c'^*) in the model being simulated. As a bootstrap estimate of a standard error is based on no analytical derivation or assumptions, it serves as a good benchmark for comparing the analytical EIV standard error estimates. As can be seen in Table 6, the average analytical standard error using expression A4 was generally smaller than the bootstrap standard error in the N=50 condition, but this difference dissipates rapidly as sample size increases. This is consistent with small-sample concerns expressed by Lockwood and McCaffrey (2020) and Counsel and Cribbie (2017). But notice that Table 6 that the standard errors generated with expressions A2/A3 that we used in our simulations and the default in PROCESS were consistently closer to the average bootstrap standard error in smaller samples, though just barely larger. This would lead to conservative tests and confidence intervals that are slightly too wide in smaller samples, though there is little evidence of this in Table 5. Nevertheless, in small samples, we suggest conducting or at least double-checking inferences for individual paths using either bootstrap confidence intervals or a bootstrap estimate of sampling variance. This can be done easily in PROCESS using the **modelbt** option. See the documentation.

It is not our intention to suggest that EIV regression is statistically better than latent variable SEM. On the contrary, we have shown that EIV regression seems to perform as well as the SILV approach. SEM offers much more flexibility in model estimation, various accepted means of dealing with missing data, measures of model fit, and other advantages beyond the scope of this paper. But we do feel that EIV regression is substantially less analytically complex than SEM methods and requires (as implemented in PROCESS) less code even for more complex structural models that can be intimidating to the average researcher to program in SEM. EIV regression as implemented in the PROCESS macro only a few additional keystrokes in the syntax compared to doing nothing and can be easily adopted by even novice researchers. Furthermore, the EIV estimation implemented in PROCESS is quite fast. PROCESS took no more than 8 seconds for each analysis reported in Tables 1-3, including the 5,000 bootstrap samples for inference about the indirect effect, whereas the SILV approach in lavaan in R using the code in Appendix B took between 60 and 90 seconds.[5]

[5] When executed on a Dell Latitude 5310 64bit Intel i5 CPU @ 1.7GHz with 8GB RAM and Windows 10 operating system. The Mplus code for the SILV approach provided in the supplementary materials ran in about 4 seconds. The corresponding Stata code required between 6 and 8 minutes to execute.

Our enthusiasm for the results we describe here is somewhat tempered by our awareness that this manuscript is more of a demonstration by example rather than a more rigorous analysis and simulation that varies widely and orthogonally various factors that may influence these results, such as the sizes of the effects being estimated, the reliability of the observed variables, sample size, and other things. We encourage future research on the performance of EIV-regression relative to alternatives to further explore the boundary conditions and generality of our claims and wisdom of our recommendations.

We conclude by warning researchers who might be excited by how easy it now is to account for random measurement error in their mediation analyses that EIV regression should not be used as a means of sweeping deeper measurement problems under the rug. As Flake and Fried (2020) discuss, many researchers are remarkably cavalier in their approach to measurement, often using questionable measurement practices that lower the quality and meaning of their data and make their results hard to interpret. Poor measurement, construct invalidity, and highly unreliable or otherwise low-quality data cannot be made righteous by statistical sleights-of-hand. The EIV implementation we have described here is no exception to this general rule. Its use is best reserved for situations in which investigators have given careful thought to how they are measuring their constructs, are convinced they are measuring what they intend to be measuring and are doing so reasonably well but find they need a little extra estimation help given their measurements still contain some inevitable random measurement error.

References

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Buonaccorsi, J. P. (2010). *Measurement error: Models, methods, and applications*. Boca Raton, FL: Chapman & Hall.

Chan, M., Hu, P., & Mak, M. K. F. (2022). Mediation analysis and warranted inferences in media and communication research: Examining research design in communication journals from 1996 to 2017. *Journalism and Mass Communication Quarterly, 99*, 463-486. DOI: https://doi.org/10.1177/107769902096151

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods, 11*, 296-325. DOI: https://doi.org/10.1177/1094428107300343

Cohen, J., Cohen, P., West, S.G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods, 19*, 300-315. DOI: https://doi.org/10.1037/a0033805

Counsell, A., & Cribbie, R. A. (2017). Using the errors-in-variables method in two-group pretest-posttest designs. *Methodology, 13*, 1-8. https://doi.org/10.1027/1614-2241/a000122

Culpepper, S. A., & Aguinis, H. (2011). Using analysis of covariance (ANCOVA) with fallible covariates. *Psychological Methods, 16*, 166-178. DOI: https://doi.org/10.1037/a0023355

Darlington, R. B., & Hayes, A. F. (2017). *Regression analysis and linear models: Concepts, application, and implementation*. New York: The Guilford Press.

Falk, C. F., & Biesanz, J. C. (2015). Inference and interval estimation methods for indirect effects with latent variable models. *Structural Equation Modeling, 22*, 24-38. https://doi.org/10.1080/10705511.2014.935266

Flake, J. K., & Fried, E. I. (2020). Measurement Schmeasurement: Questionable measurement practices and how to avoid then. *Advances in Methods and Practices in Psychological Science, 3*, 456-465. https://doi.org/10.1177/2515245920952393

Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. *Multivariate Behavioral Research, 51*, 681-697. https://doi.org/10.1080/00273171.2016.1224154

Fuller, W. A. (1987). *Measurement error models*. New York: Wiley.

Fye, H. J., Kim, Y., & Rainey, J. S. (2022). Correlates of wellness, affective distress, and burnout among school counselors. *Journal of Counseling Development, 100*, 308-318. https://doi.org/10.1002/jcad.12493

Gaboury, M. P., Belleville, G., Lebel, J. et al. (2023). Posttraumatic growth and its relationship with posttraumatic stress symptoms following wildfires: A longitudinal study. *Traumatology*. https://doi.org/10.1037/trm0000470

Gainsburg, I., & Cunningham, J. L. (2023). Compassion fatigue as a self-fulfilling prophecy: Believing compassion is limited increases fatigue and decreases compassion. *Psychological Science, 34*(11), 1206-1219. https://doi/10.1177/09567976231194537

Gonzales, O., & MacKinnon, D. P. (2021). The measurement of the mediator and its influence on statistical mediation conclusions. *Psychological Methods, 26*, 1-17. DOI: https://doi.org/10.1037/met0000263

Grisbook, M.-A., Dewey, D., Cuthbert, C., McDonald, S., Ntanda, H., & Letourneau, N. (2024). The association between Cesarean section delivery and child behavior: Is it mediated by maternal post-traumatic stress disorder and maternal postpartum depression? *Behavioral Sciences, 14*, 61. https://doi.org/10.3390/bs14010061

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76, 408-420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression based approach* (3rd Ed.). New York, NY: The Guilford Press.

Hayes, A. F., & Cai, L. (2007). Using heteroscedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods, 39*, 709-722.

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal, 25*, 76-81. https://doi.org/10.1016/j.ausmj.2017.02.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, 67, 451-470. https://doi.org/10.1111/bmsp.12028

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science, 24*,1918-1927. https://doi.org/10.1177/0956797613480187

Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science, 25*, 354-359. https://doi.org/10.1177/0956797613502676

Kline, R. B. (2023). Principles and practice of structural equation modeling (5th Edition). New York: The Guilford Press.

Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods, 15*, 3-16. https://doi.org/10.1177/1094428110391673

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*, 1174-1188. https://doi.org/10.1037/a0024776

Lockwood, J. R., & McCaffrey, D. F. (2020). Recommendations about estimating errors-in-variables regression in Stata. *The Stata Journal, 20*, 116-130.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity-consistent standard errors in the linear regression model. *American Statistician, 54*, 217-224.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. Psychological Methods, 7, 19-40. https://doi.org/10.1037/1082-989X.7.1.19

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

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*, 305-325.

Neufeld, A., & Malin, G. (2022). Need fulfillment and resilience mediate the relationship between mindfulness and coping in medical students. *Teaching and Learning in Medicine, 34*, 78-88.

Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs, 79*, 171-181. https://doi.org/10.15288/jsad.2018.79.171

Ozimek, P., Lainas, S., Bierhoff, H.-W., & Rohmann, E. (2023). How photo editing in social media shapes self-perceived attractiveness and self-esteem via self-objectification and physical appearance comparisons. *BMC Psychology, 11*(99), 1-14. https://doi.org/10.1186/s40359-023-01143-0

Pieters, Rik (2017). Meaningful mediation analysis: Plausible causal inference and informative communication. *Journal of Consumer Research*, *44*, 692–716

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers, 36*, 717-731.

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

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

Rokeach, A., & Wiener, J. (2022). Predictors of friendship quality in adolescents with and without attention-deficit/hyperactivity disorder. *School Mental Health, 14*, 328-340. https://doi.org/10.1007/s12310-022-09508-3

Shear, B. R., & Zumbo, B. D. (2013). False positives in multiple regression: Unanticipated consequences of measurement error in the predictor variables. *Educational and Psychological Measurement, 73*, 733-756. https://doi.org/10.1177/0013164413487738

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422-445. https://doi.org/10.1037/1082-989X.7.4.422

Smith, A., Andruski, B., Deng, G., & Burnham, R. (2022). The association between sleep disturbance and health outcomes in chronic whiplash-associated disorders: A series of mediation analyses. *The Clinical Journal of Pain, 38,* 612-619. https://doi.org/10.1097/AJP.0000000000001066

StataCorp (2023). *Stata 18 Base Reference Manual*. College Station, TX: Stata Press.

Tingley, D., Yamamoto, T., Hirose, K., Imai, K. & Keele, L. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59* (5), 1-38. https://doi.org/10.18637/jss.v059.i05

Valeri, L., Lin, X., & VanderWeele, T. J. (2014). Mediation analysis when a continuous mediator is measured with error and the outcome follows a generalized linear model. *Statistics in Medicine, 33*, 4875-4890. https://doi.org/10.1002/sim.6295

VanderWeele, T. J., Valeri, L., & Ogburn, E. L. (2012). The role of measurement error and misclassification in mediation analysis. *Epidemiology, 23*, 561-564. https://doi.org/10.1097/EDE.0b013e318258f5e4

Wang, P., Sang, J., Li, P. & Zhao, J. (2016). How to make a newcomer happy? The mediating role of career commitment on the relationship between unmet expectations and job satisfaction. Social Indicators Research, 127, 401-412. https://doi.org/10.1007/s11205-015-0952-3

Yang, Y., Sedikides, C., Wang, Y., & Cai, H. (2024). Nature nurtures authenticity: Mechanisms and consequences. *Journal of Personality and Social Psychology, 126*(1), 79-104. https://doi.org/10.1037/pspi0000432

Yasuda, Y., & Goegan, L. D. (2023). Effective strategy types depending on regulatory focus in academic settings. *Learning and Motivation, 83*, 101906. https://doi.org/10.1016/j.lmot.2023.101906

Table 1.

A mediation analysis of the effect of compassion mindset on compassion fatigue through expected fatigue from an experimental task using ordinary least squares regression assuming perfect reliability (OLS), a single-indicator latent variable model (SILV), and errors-in-variables regression (EIV) as implemented in PROCESS. Standard errors in parentheses. Standard errors for the indirect effect are bootstrap estimates. $n = 308$.

	Model of Expected Fatigue (*M*)			Model of Compassion Fatigue (*Y*)		
	OLS	EIV	SILV	OLS	EIV	SILV
Constant	4.266 (0.210)	4.630 (0.288)	---	3.004 (0.302)	2.916 (0.464)	---
Compassion Mindset (*X*)	−0.250 (0.052)	−0.342 (0.075)	−0.343 (0.072)	−0.146 (0.050)	−0.178 (0.071)	−0.178 (0.072)
Expected Fatigue (*M*)	---	---	---	0.464 (0.054)	0.529 (0.084)	0.529 (0.066)
R^2	.071	.097		.254	.293	

	OLS	EIV	SILV
Total effect of Compassion Mindset	−0.262 (0.054)	−0.359 (0.078)	−0.359 (0.075)
Direct effect of Compassion Mindset	−0.146 (0.050)	−0.178 (0.071)	−0.178 (0.072)
Indirect effect of Compassion Mindset	−0.116 (0.032)	−0.181 (0.051)	−0.181 (0.050)
[95% percentile bootstrap CI]	[−0.185, −0.059]	[−0.292, −0.093]	[−0.292, −0.093]

Note: SILV constants are not provided because it estimates a different quantity than OLS or EIV constants given the arbitrary scaling of latent variables to have mean of zero.

Table 2.

A mediation analysis of the effect of experimental nature exposure on self-authenticity through positive affect using ordinary least squares regression assuming perfect reliability (OLS), a single-indicator latent variable model (SILV), and errors-in-variables regression (EIV) as implemented in PROCESS. Standard errors in parentheses. Standard errors for the indirect effect are bootstrap estimates. $n = 171$

	Model of Positive Affect (*M*)			Model of Authenticity (*Y*)		
	OLS	EIV	SILV	OLS	EIV	SILV
Constant	3.433 (0.063)	3.433 (0.068)	---	2.823 (0.392)	2.651 (0.471)	---
Nature Condition (*X*)	0.320 (0.089)	0.320 (0.090)	0.320 (0.089)	0.143 (0.133)	0.127 (0.139)	0.127 (0.133)
Positive Affect (*M*)	---	---	---	0.468 (0.111)	0.519 (0.133)	0.519 (0.122)
R^2	.070	.070		.121	.131	

	OLS	EIV	SILV
Total effect of Nature Condition	0.292 (0.135)	0.292 (0.136)	0.292 (0.134)
Direct effect of Nature Condition	0.143 (0.133)	0.127 (0.139)	0.127 (0.133)
Indirect effect of Nature Condition	0.150 (0.053)	0.166 (0.059)	0.166 (0.058)
[95% percentile bootstrap CI]	[0.058, 0.263]	[0.064, 0.293]	[0.063, 0.289]

Note: SILV constants are not provided because they estimate a different quantity than OLS or EIV constants given the arbitrary scaling of latent variables to have mean of zero.

Table 3.

A mediation analysis of the effect of photo-editing behaviours on self-perceived attractiveness through self-objectifying using ordinary least squares regression assuming perfect reliability (OLS), a single-indicator latent variable model (SILV), and errors-in-variables regression (EIV) as implemented in PROCESS. Standard errors in parentheses. Standard errors for the indirect effect are bootstrap estimates. $n = 671$

	Model of Self-Objectifying (*M*)			Model of Self-Perceived Attractiveness (*Y*)		
	OLS	EIV	SILV	OLS	EIV	SILV
Constant	1.544 (0.073)	1.283 (0.104)	---	1.551 (0.042)	1.453 (0.061)	---
Photo Editing (*X*)	0.405 (0.035)	0.540 (0.051)	0.540 (0.048)	0.082 (0.017)	0.090 (0.029)	0.090 (0.026)
Self-objectifying (*M*)	---	---	---	0.317 (0.017)	0.353 (0.023)	0.353 (0.022)
R^2	.166	.222		.434	.482	

	OLS	EIV	SILV
Total effect of Photo Editing	0.210 (0.019)	0.280 (0.032)	0.280 (0.026)
Direct effect of Photo Editing	0.082 (0.017)	0.090 (0.029)	0.090 (0.026)
Indirect effect of Photo Editing	0.129 (0.014)	0.191 (0.022)	0.191 (0.021)
[95% percentile bootstrap CI]	[0.101, 0.157]	[0.149, 0.234]	[0.152, 0.232]

Note: SILV constants are not provided because they estimate a different quantity than OLS or EIV constants given the arbitrary scaling of latent variables to have mean of zero.

Table 4.

Mean Estimated Effect (“Mean”) and Mean Bias Percentage (“Bias%”) in the Simulation (True Effect Being Estimated in Parentheses).

Compassion Fatigue

	Indirect Effect (–0.116)				Direct Effect (–0.146)				Total Effect (–0.262)			
n	OLS		EIV		OLS		EIV		OLS		EIV	
	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%
50	-.074	-36.6	-.115	-0.8	-.117	-19.8	-.146	0.1	-.191	-27.2	-.261	-0.3
100	-.075	-35.6	-.117	0.9	-.118	-19.3	-.147	0.5	-.192	-26.5	-.264	0.6
200	-.073	-36.7	-.115	-1.0	-.118	-19.5	-.147	0.4	-.191	-27.1	-.261	-0.2
300	-.074	-36.2	-.116	-0.1	-.117	-19.7	-.146	0.1	-.191	-27.0	-.262	0.0
500	-.074	-35.9	-.116	0.3	-.118	-19.2	-.147	0.8	-.192	-26.6	-.264	0.6
1000	-.074	-35.9	-.116	0.2	-.118	-19.3	-.147	0.6	-.192	-26.7	-.263	0.4

Nature and Self-Actualization

	Indirect Effect (0.150)				Direct Effect (0.143)				Total Effect (0.293)			
n	OLS		EIV		OLS		EIV		OLS		EIV	
	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%
50	.135	-10.1	.150	0.2	.153	6.7	.137	-4.1	.287	-1.9	.287	-1.9
100	.137	-8.5	.152	1.5	.157	10.0	.142	-0.6	.294	.5	.294	0.5
200	.135	-10.1	.149	-0.4	.158	10.3	.143	0.1	.292	-.1	.292	-0.1
300	.135	-9.6	.150	0.1	.160	12.1	.146	1.9	.296	1.0	.296	1.0
500	.135	-10.2	.149	-0.6	.159	10.9	.144	0.8	.293	.1	.293	0.1
1000	.136	-9.5	.150	0.2	.157	9.9	.143	-0.3	.293	.0	.293	0.0

Photo Editing

	Indirect Effect (0.128)				Direct Effect (0.082)				Total Effect (0.210)			
n	OLS		EIV		OLS		EIV		OLS		EIV	
	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%	Mean	Bias%
50	.087	-32.2	.129	0.5	.070	-14.4	.081	-1.7	.157	-25.2	.210	-0.3
100	.087	-32.2	.129	0.3	.070	-14.3	.081	-1.3	.157	-25.2	.210	-0.3
200	.087	-32.6	.128	-0.4	.071	-13.2	.083	0.6	.158	-25.0	.210	0.0
300	.087	-32.3	.128	0.0	.071	-13.4	.082	0.1	.158	-25.0	.211	0.0
500	.087	-32.1	.129	0.4	.071	-14.0	.081	-0.7	.158	-25.0	.210	0.0
1000	.087	-32.5	.128	-0.2	.071	-13.8	.082	-0.4	.157	-25.2	.210	-0.3

Table 5

Percentage of OLS and EIV (from expressions A2 and A3) 95% Confidence Intervals in the Simulation Containing the True Value of the Effect Being Estimated.

Compassion Fatigue

n	Indirect Effect		Direct Effect		Total Effect	
	OLS	EIV	OLS	EIV	OLS	EIV
50	84.3	94.3	93.8	94.6	91.4	94.0
100	79.7	94.7	93.7	94.9	88.1	94.4
200	66.9	94.5	92.0	94.5	80.0	94.8
300	55.6	95.2	90.8	95.1	72.4	95.0
500	38.0	94.7	89.2	95.0	58.5	95.0
1000	11.7	94.7	82.0	94.7	31.0	95.4

Nature and Self-Actualization

n	Indirect Effect		Direct Effect		Total Effect	
	OLS	EIV	OLS	EIV	OLS	EIV
50	92.1	93.5	95.5	96.2	95.0	95.6
100	93.7	94.8	95.3	95.7	94.8	94.9
200	92.6	94.7	95.3	95.5	94.9	95.0
300	92.6	94.7	94.5	94.4	94.5	94.6
500	91.9	95.1	94.3	94.7	94.7	94.8
1000	89.4	95.5	94.6	95.8	95.8	95.8

Photo Editing

n	Indirect Effect		Direct Effect		Total Effect	
	OLS	EIV	OLS	EIV	OLS	EIV
50	79.7	94.4	94.7	95.2	87.1	94.4
100	69.8	94.8	94.5	95.4	79.8	95.3
200	48.1	94.6	93.6	95.1	62.8	94.9
300	32.1	94.6	92.3	95.4	50.1	95.3
500	13.5	94.9	90.5	94.9	27.7	94.8
1000	00.9	95.4	85.5	94.8	04.8	94.8

Table 6

Mean (over 5,000 simulation repetitions) Analytical (from Appendix A, expressions A2 and A3; or A4) and Bootstrap Estimates of EIV Standard Errors of the Paths in the Simulation. Table Entries Under “Boot” are the Average Standard Deviation of the 5,000 Bootstrap Estimates of the Path.

Compassion Fatigue

n	a^*			b^*			c'^*		
	A2/A3	A4	Boot	A2/A3	A4	Boot	A2/A3	A4	Boot
50	.173	.160	.168	.182	.165	.177	.185	.167	.182
100	.118	.114	.117	.123	.117	.119	.125	.119	.125
200	.083	.081	.081	.085	.083	.085	.086	.084	.087
300	.067	.067	.068	.069	.068	.068	.070	.069	.069
500	.052	.052	.052	.053	.052	.052	.054	.053	.053
1000	.037	.037	.037	.037	.037	.037	.038	.038	.038

Nature and Self-Actualization

n	a^*			b^*			c'^*		
	A2/A3	A4	Boot	A2/A3	A4	Boot	A2/A3	A4	Boot
50	.176	.169	.174	.256	.232	.253	.287	.269	.281
100	.124	.121	.122	.175	.166	.173	.199	.193	.196
200	.087	.086	.088	.122	.119	.120	.139	.137	.137
300	.071	.071	.071	.099	.097	.099	.113	.112	.114
500	.055	.055	.054	.076	.075	.076	.088	.087	.088
1000	.039	.039	.038	.054	.053	.054	.062	.062	.061

Photo Editing

n	a^*			b^*			c'^*		
	A2/A3	A4	Boot	A2/A3	A4	Boot	A2/A3	A4	Boot
50	.171	.158	.167	.084	.076	.082	.091	.082	.088
100	.118	.113	.115	.057	.054	.056	.061	.058	.059
200	.082	.080	.081	.039	.038	.039	.042	.041	.041
300	.067	.066	.066	.032	.031	.032	.034	.034	.034
500	.052	.051	.051	.024	.024	.025	.026	.026	.026
1000	.036	.036	.036	.017	.017	.017	.019	.018	.019

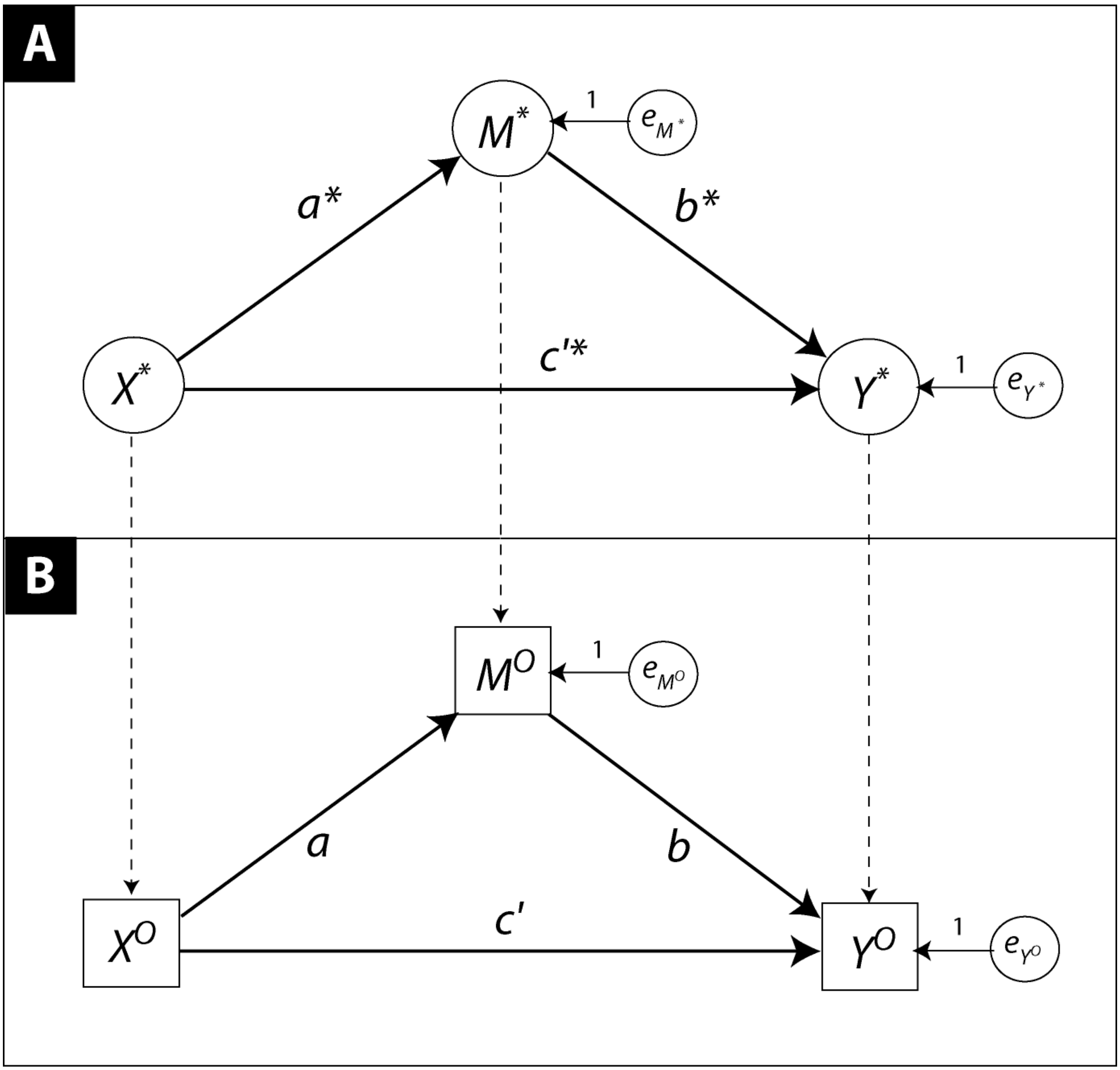

Figure 1. A mediation model researchers typically want to estimate (panel A) versus what they typically estimate (panel B). See equations 1 through 4. Dashed arrows reflect the measurement assumption that true scores (superscripts) affect observed scores (O superscripts).*

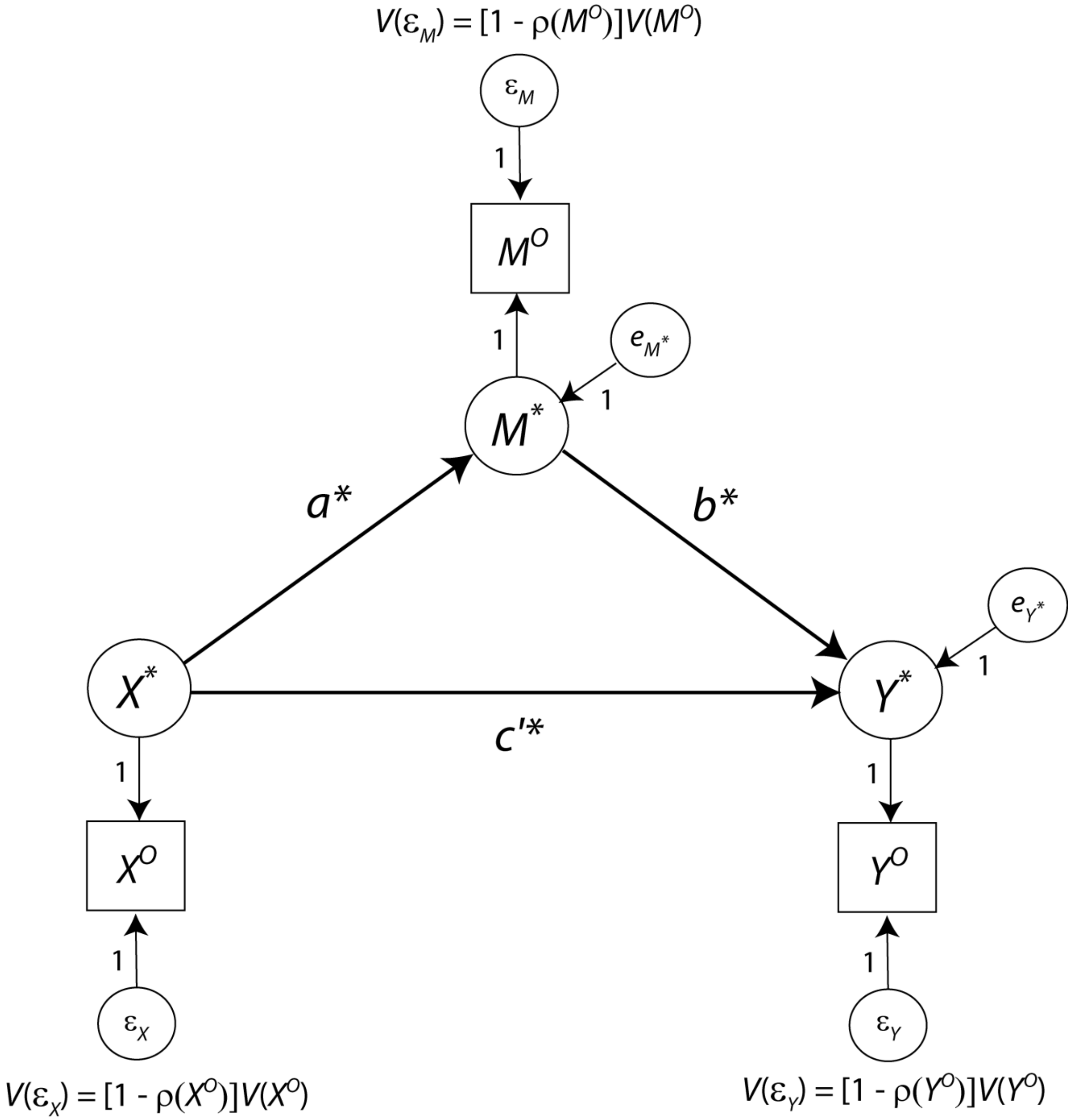

*Figure 2. A single indicator latent variable (SILV) mediation model with one mediator in the notation used throughout this paper. Variables with * superscripts are latent variables. Variables with O superscripts are observed variables. V = variance.*

Appendix A

Errors-in-Variables Regression Computations

The errors-in-variables regression analysis algorithms we use in our examples, the simulation, and that are implemented in PROCESS were motivated by computations done by Stata as of version 15 and described in StataCorp (2023, pp. 594-600), which we report below for convenience and using similar notation.

Let **X** be an $n \times (k+1)$ matrix containing the observed data from *n* observations for *k* variables on the right-hand side of a regression model, with the last column containing all ones for the regression constant. Let **y** be an $n \times 1$ vector of observed measurements of *Y*, the variable on the left side of the regression equation. And let **E** be a $(k+1)$ diagonal matrix with the *j*th diagonal element set to $n[1-\rho(x_j)]V(x_j)$ where $\rho(x_j)$ and $V(x_j)$ are the reliability and variance, respectively, of the observed data in the *j*th column of **X**. And, let $\mathbf{D} = \mathbf{X}'\mathbf{X} - \mathbf{E}$. The elements of $\mathbf{b} = \mathbf{D}^{-1}\mathbf{X}'\mathbf{y}$ are the EIV estimates of the regression coefficients for the *k* variables in **X** in the model of *Y*, with the last entry being the regression constant. The mean squared error for the EIV model is $s^2 = (\mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{D}^{-1}\mathbf{b})/(n-k-1)$ and the squared multiple correlation is $1 - [(n-k-1)s^2/SS_{total})]$ where SS_{total} is the total sum of squares calculated in the usual way in regression analysis.

The variance-covariance matrix of **b** is needed for inference about regression coefficients. Define

$$e_i x_{ij} + (x_{ij} - \bar{x}_j)^2[1-\rho(x_j)]b_j \qquad \text{(A1)}$$

as the element in the *i*th row and *j*th column of an $n \times (k+1)$ matrix **H** where *i* is the observation in the *i*th row in **X**, *j* is the variable in the *j*th column of **X**, and $e_i = y_i - x_i\mathbf{b}$. The variance-covariance matrix of **b** is estimated as

$$\mathbf{D}^{-1}\mathbf{H}'\mathbf{H}\mathbf{D}^{-1} \qquad \text{(A2)}$$

(Stefanski & Boos, 2002, Buonaccorsi, 2010, Fuller, 1987, as cited in StataCorp, 2023) and the square root of the diagonal elements in this resulting $(k+1) \times (k+1)$ matrix are the standard errors of the regression coefficients in **b**. Note that the standard errors estimated in this manner do *not* converge to the OLS standard errors as the reliability of observed data in columns of **X** converge to 1. This is because this approach to EIV standard error estimation includes a robustification component to offset the effects of heteroscedasticity of unknown form. When reliabilities are all set to 1, expression A1 generates heteroscedasticity-consistent standard errors equivalent to HC0 described in Long and Ervin (2000) and Hayes and Cai (2007). Regardless of the setting for reliabilities in the EIV routine, this method, when requested using option **eiv=0** in the PROCESS command, will produce standard errors equivalent to those generated by Stata as of version 15.

But the HC0 correction for heteroscedasticity can produce standard errors that are too small in smaller samples (see Long & Ervin, 2000), resulting in elevated Type I errors and lower confidence interval coverage than the nominal level. Our (unpublished) simulation results examining the performance of expressions A1 and A2 are consistent with this. PROCESS can implement a different EIV variance-covariance matrix estimator for **b** with an alternative heteroscedasticity correction that yields standard errors equivalent to HC3 (MacKinnon and White, 1985) when all reliabilities are set to 1. Research shows in the perfect reliability case that the HC3 estimator performs better in smaller samples than does HC0 (Long & Ervin, 2000). Defining h_i as case *i*'s leverage, which is the *i*th diagonal element of $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}$, the *i*th row and *j*th column of **H** are set to

$$(e_i x_{ij})/(1 - h_i) + (x_{ij} - \bar{x}_j)^2[1 - \rho(x_j)]b_j \qquad \text{(A3)}$$

Expression A2 then generates the variance-covariance matrix of **b**, the diagonals of which are the squared standard errors for the regression coefficients and constant. This is the default variance-covariance estimator of **b** used by PROCESS's EIV regression routine (or is specified using **eiv=3** in the PROCESS command) and this is the standard error estimator that was used in the example analyses and simulations we report in this manuscript. This approach is not available in Stata or any other software other than PROCESS as of the writing of this manuscript.

PROCESS can also compute a third and somewhat simpler approach to constructing the variance-covariance matrix of **b**, implemented in Stata prior to version 15 and described in Lockwood and McCaffrey (2000). By including **eiv=5** in the PROCESS command, the standard errors for the EIV regression coefficients are calculated as the square root of the diagonal elements of

$$s^2\mathbf{D}^{-1}\mathbf{X}'\mathbf{X}\mathbf{D}^{-1}. \tag{A4}$$

This approach to estimating the standard error of the regression coefficients will be identical to the OLS standard errors when the reliabilities of all the variables on the right-hand side of the equation are equal to 1. However, Lockwood and McCaffrey (2000) show analytically as well as through simulation that this approach can produce standard errors that are too small in some circumstances. This approach to estimating the variance-covariance matrix of **b** was not used in our simulations or example computations, and it does not account for heteroscedasticity of unknown form as the other methods implemented in PROCESS do.

Appendix B

Code for the Three Example Mediation Analyses

This appendix provides the code to conduct the analyses reported in Tables 1-3. The data are available at https://osf.io/x8we5/.

Compassion Fatigue and Compassion Mindset

Observed Variable OLS Regression Using PROCESS.

SPSS: `process y=fatigue/x=compass/m=efatigue/model=4/total=1/seed=24080.`

SAS: `%process(data=compfat,y=fatigue,x=compass,m=efatigue,model=4,total=1, seed=24080)`

R: `process(data=compfat,y="fatigue",x="compass",m="efatigue",model=4,total=1, seed=24080)`

Errors-in-Variables Regression Using PROCESS (requires PROCESS version 5 or later)

SPSS: `process y=fatigue/x=compass/m=efatigue/model=4/total=1/seed=24080 /relx=0.73/relm=0.86.`

SAS: `%process(data=compfat,y=fatigue,x=compass,m=efatigue,model=4,total=1, seed=24080,relx=0.73,relm=0.86)`

R: `process(data=compfat,y="fatigue",x="compass",m="efatigue",model=4,total=1, seed=24080,relx=0.73,relm=0.86)`

Single-Indicator Latent Variable Structural Equation Model in R using lavaan

```
compfat<-read.table("compfat.csv", sep=",",header=TRUE)
library(lavaan)
model.silv<-"Lfatigue=~fatigue
             Lcompass=~compass
             Lefatigue=~efatigue
             Lefatigue~a*Lcompass
             Lfatigue~b*Lefatigue+cp*Lcompass
             ab := a*b
              c :=a*b+cp
             #(1-reliability) multiplied by observed variances
             compass~~((1-0.73)*0.912239)*compass
             fatigue~~((1-0.84)*0.879927)*fatigue
             efatigue~~((1-0.86)*0.804416)*efatigue"
modelp<-sem(model.silv,data=compfat)
summary(modelp,fit.measures=TRUE)
```

```
set.seed(24080)
modelp<-sem(model.silv,data=compfat,se="bootstrap",bootstrap=5000)
parameterestimates(modelp,boot.ci.type="perc")
```

Nature and Self-Actualization

Observed Variable OLS Regression Using PROCESS.

SPSS: `process y=authcity/x=natcond/m=paffect/model=4/total=1/seed=27654.`

SAS: `%process (data=nature,y=authcity,x=natcond,m=paffect,model=4,total=1, seed=27654)`

R: `process(nature,y="authcity",x="natcond",m="paffect",model=4,total=1, seed=27654)`

Errors-in-Variables Regression using PROCESS (requires PROCESS version 5 or later)

SPSS: `process y=authcity/x=natcond/m=paffect/model=4/total=1/seed=27654 /relm=0.91.`

SAS: `%process (data=nature,y=authcity,x=natcond,m=paffect,model=4,total=1, seed=27654,relm=0.91)`

R: `process(nature,y="authcity",x="natcond",m="paffect",model=4,total=1, seed=27654,relm=0.91)`

Single-Indicator Latent Variable Structural Equation Model in R using lavaan

```
nature<-read.table("nature.csv", sep=",",header=TRUE)
library(lavaan)
model.silv<-"Lauthcity=~authcity
             Lpaffect=~paffect
             Lpaffect~a*natcond
             Lauthcity~b*Lpaffect+cp*natcond
             ab :=a*b
             c := a*b+cp
             #(1-reliability) multiplied by observed variances
             authcity~~((1-0.82)*0.7950942)*authcity
             paffect~~((1-0.91)*0.3656030)*paffect"
modelp<-sem(model.silv,data=nature)
summary(modelp,fit.measures=TRUE)
set.seed(27654)
modelp<-sem(model.silv,data=nature,se="bootstrap",bootstrap=5000)
parameterestimates(modelp,boot.ci.type="perc")
```

Photo Editing and Self-Perceived Attractiveness

Observed Variable OLS Regression Using PROCESS.

SPSS: `process y=spa/x=pes/m=sobbs/model=4/total=1/seed=7234.`

SAS: `%process (data=photo,y=spa,x=pes,m=sobbs,model=4,total=1,seed=7234)`

R: `process(photo,y="spa",x="pes",m="sobbs",model=4,total=1,seed=7234)`

Errors-in-Variables Regression using PROCESS (requires PROCESS version 5 or later)

SPSS:
```
process y=spa/x=pes/m=sobbs/model=4/total=1/seed=7234/relx=0.75
   /relm=0.89.
```

SAS:
```
%process (data=photo,y=spa,x=pes,m=sobbs,model=4,total=1,seed=7234,
   relx=0.75,relm=0.89)
```

R:
```
process(photo,y="spa",x="pes",m="sobbs",model=4,total=1,seed=7234,
   relx=0.75,relm=0.89)
```

Single-Indicator Latent Variable Structural Equation Model in R using lavaan

```
photo<-read.table("photo.csv", sep=",",header=TRUE)
library(lavaan)
model.silv<-"Lspa=~spa
             Lpes=~pes
             Lsobbs=~sobbs
             Lsobbs~a*Lpes
             Lspa~b*Lsobbs+cp*Lpes
             ab :=a*b
             c := a*b+cp
             #(1-reliability) multiplied by observed variances
             spa~~((1-0.94)*0.1602310)*spa
             pes~~((1-0.75)*0.5474906)*pes
             sobbs~~((1-0.89)*0.5406272)*sobbs"
modelp<-sem(model.silv,data=nature)
summary(modelp,fit.measures=TRUE)
set.seed(7234)
modelp<-sem(model.silv,data=photo,se="bootstrap",bootstrap=5000)
parameterestimates(modelp,boot.ci.type="perc")
```

Appendix C

PROCESS Output from a Mediation Analysis using EIV Regression

The PROCESS output below was generated with the PROCESS command provided in Appendix B.

```
Run MATRIX procedure:

***************** PROCESS Procedure for SPSS Version 5.0 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
  Model: 4
      Y: fatigue
      X: compass
      M: efatigue

Sample
Size:  308

Custom
Seed:     24080

**************************************************************************
                     Errors-in-variables regression

OUTCOME VARIABLE:
 efatigue

Model Summary
       R-sq        MSE     F(HC3)        df1        df2          p
      .0969      .7289    20.9460     1.0000   306.0000      .0000

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     4.6305      .2875    16.1057      .0000     4.0647     5.1962
compass      -.3421      .0747    -4.5767      .0000     -.4891     -.1950

**************************************************************************
                     Errors-in-variables regression

OUTCOME VARIABLE:
 fatigue

Model Summary
       R-sq        MSE     F(HC3)        df1        df2          p
      .2926      .6265    31.0888     2.0000   305.0000      .0000

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     2.9162      .4640     6.2850      .0000     2.0031     3.8292
compass      -.1779      .0711    -2.5012      .0129     -.3178     -.0379
efatigue      .5290      .0838     6.3144      .0000      .3641      .6938
```

```
************************* TOTAL EFFECT MODEL ***************************
                     Errors-in-variables regression

OUTCOME VARIABLE:
 fatigue

Model Summary
       R-sq        MSE     F(HC3)        df1        df2          p
      .0974      .7968    21.3301     1.0000   306.0000      .0000

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     5.3655      .3035    17.6803      .0000     4.7683     5.9627
compass      -.3588      .0777    -4.6185      .0000     -.5117     -.2059


************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y
     Effect    se(HC3)          t          p       LLCI       ULCI
     -.3588      .0777    -4.6185      .0000     -.5117     -.2059

Direct effect of X on Y
     Effect    se(HC3)          t          p       LLCI       ULCI
     -.1779      .0711    -2.5012      .0129     -.3178     -.0379

Indirect effect(s) of X on Y:
             Effect     BootSE   BootLLCI   BootULCI
efatigue     -.1809      .0514     -.2922     -.0926

********************** ANALYSIS NOTES AND ERRORS ***********************

Level of confidence for all confidence intervals in output:
  95.0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000

NOTE: A heteroskedasticity-consistent standard error and covariance matrix estimator
was used.

NOTE: This errors-in-variables analysis assumes the following reliabilities:
  compass efatigue
    .7300    .8600
```