SAS® Macros for Computing Causal Mediated Effects in Two- and Three-Wave Longitudinal Models

Matthew J. Valente, David P. MacKinnon, Arizona State University

ABSTRACT

Mediation analysis is a statistical technique for investigating the extent to which a mediating variable transmits the effect of an independent variable to a dependent variable. Because it is used in many fields, there have been rapid developments in statistical mediation. The most cutting-edge statistical mediation analysis focuses on the causal interpretation of mediated effects. Causal inference is particularly challenging in mediation analysis because of the difficulty of randomizing subjects to levels of the mediator. The focus of this paper is on updating three existing SAS® macros (%TWOWAVEMED, %TWOWAVEMONTECARLO, and %TWOWAVEPOSTPOWER, presented at SAS® Global Forum 2017) in two important ways. First, the macros are updated to incorporate new cutting-edge methods for estimating longitudinal mediated effects from the Potential Outcomes Framework for causal inference. The two new methods are inverse-propensity weighting, an application of propensity scores, and sequential G-estimation. The causal inference methods are revolutionary because they frame the estimation of mediated effects in terms of differences in potential outcomes, which align more naturally with how researchers think about causal inference. Second, the macros are updated to estimate mediated effects across three waves of data. The combination of these new causal inference methods and three waves of data enable researchers to test how causal mediated effects develop and maintain over time.

INTRODUCTION

Statistical mediation analysis is a common statistical technique used in marketing research, social sciences, epidemiology, and other related fields because it allows researchers to investigate *how* and through what *mechanism* two variables are related (MacKinnon, 2008). For example, mediation analysis was used to determine that self-regulatory depletion (e.g., feeling tired) leads to an increased probability of purchasing a product through its effect on increasing product-attitude certainty (i.e., how certain consumers either liked or disliked a product) (Wan, Rucker, Tormala, & Clarkson, 2010). Mediation analysis is also central to the understanding of the effects of interventions to reduce substance use behavior (Larimer et al., 2007; MacKinnon et al., 2001; West & Aiken, 1997; Witkiewitz & Bowen, 2010). Despite the advantage of investigating mechanisms of theoretical relations, it is difficult to make cause-and-effect inferences when using mediation analysis without either the use of longitudinal measurements of the mediating and outcome variables, or adjustment for potential common causes of the mediating and outcome variables (i.e., confounders), or both (Bullock, Green, & Ha, 2010; Imai, Keele, Tingley, & Yamamoto, 2011; MacKinnon, 2008; MacKinnon & Pirlott, 2015; Robins & Greenland, 1992). This paper focuses on statistical methods for mediation analysis that make use of both longitudinal measurements of the mediating and outcome variables as well as applies revolutionary causal inference methods for confounder adjustment from the potential outcomes framework. The causal inference methods are revolutionary because they frame the estimation of mediated effects in terms of differences in potential outcomes, which align more naturally with how researchers think about causal inference. This paper focuses on the widely-used randomized groups research design with pre-test and multiple post-tests that allows researchers to use longitudinal data and cutting-edge potential outcomes framework methods to make cause-and-effect inferences in a mediation analysis.

SINGLE MEDIATOR MODEL

Statistical mediation is a type of analysis that allows researchers to test indirect effects of an independent variable on a dependent variable through the independent variable's effect on the mediating variable (Lazarsfeld, 1955; MacKinnon, 2008; MacKinnon & Dwyer, 1993; Sobel 1990). A mediating variable is a type of third variable that can explain the relation between two variables. When the third variable is

conceptualized in a causal sequence between two other variables, it is called a mediator (James & Brett, 1984; MacKinnon, 2008). Statistical mediation is represented by three linear regression equations. Equation 1 represents the total effect of X on Y (c coefficient), Equation 2 represents the effect of X on M (a coefficient), and Equation 3 represents the effect of X on Y adjusted for M (c' coefficient) and the effect of M on Y adjusted for X (b coefficient). Computing the product of a and b linear regression coefficients from Equation 2 and Equation 3, respectively, represents the mediated effect of X on Y through M (ab) (See Figure 1A).

$$Y = i_1 + cX + e_1 \tag{1}$$
$$M = i_2 + aX + e_2 \tag{2}$$
$$Y = i_3 + c'X + bM + e_3 \tag{3}$$

The mediated effect is not generally interpretable as a causal effect because even if subjects are randomized to levels of the independent variable, X, subjects are not simultaneously randomized to levels of the mediator, M, therefore the mediator to outcome relation may be biased because of unmeasured confounders (Bullock, Green, & Ha, 2010; Imai, et al., 2011; MacKinnon, 2008; MacKinnon & Pirlott, 2015; Robins & Greenland, 1992). The potential outcomes framework for causal inference (Holland, 1986, 1988, Rubin, 1974) has clarified the assumptions necessary for causal interpretation of mediated effects (Robins & Greenland, 1992; Imai, Keele, Tingley, 2010; Pearl 2001, 2009, 2014; VanderWeele & Vansteelandt, 2009). One way to strengthen the causal inference regarding mediating processes is through longitudinal studies of the mediator – outcome relation which provide evidence of temporal precedence of the mediator to outcome relation (i.e., does M precede Y or does Y precede M, in time) (Cole & Maxwell, 2003; MacKinnon, 1994, 2008; Maxwell & Cole, 2007; Maxwell, Cole, & Mitchell 2011) and application of methods from the potential outcomes framework which have provided important insights into confounder adjustment for causal mediated effect estimation. A simple longitudinal experimental design that can help provide evidence of temporal precedence and be used in conjunction with confounder adjustment methods from the potential outcomes framework is the pretest-posttest control group design with multiple posttest follow-up waves.
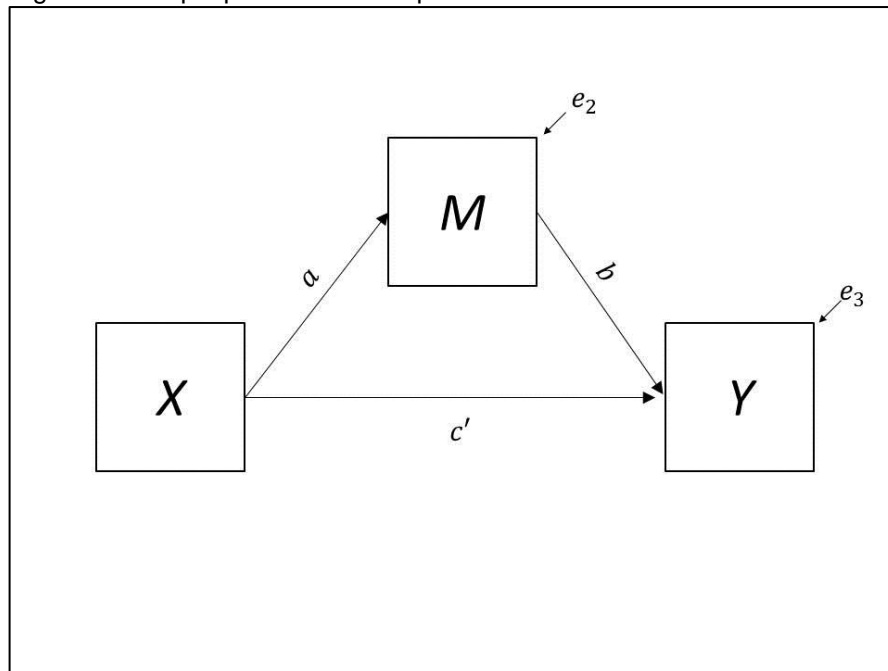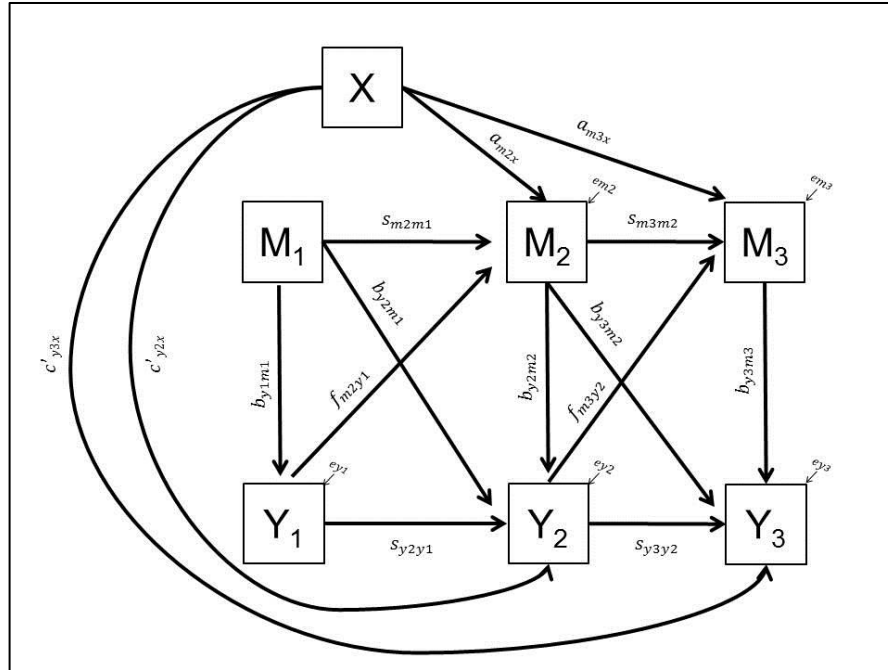

Figure 1.
Single Mediator model

THREE-WAVE LONGITUDINAL MEDIATION MODEL

Assuming X represents an experimental manipulation in a randomized experiment, evidence of the temporal order M and Y can be established with repeated measurements of M and Y over time, or longitudinal data (Gollob & Reichardt, 1991). Data is collected on M and Y before units are randomized to

levels of X which we will call pretest measures (M_1 and Y_1). Units are randomized to levels of X and then subsequent measures of the M and Y are collected which we will call posttest measures (M_2 and Y_2 etc.). To determine whether M precedes Y or Y precedes M, at least two posttest measures of M and Y are needed assuming correct timing of the mediating process has been observed (Cole & Maxwell, 2003; Collins & Graham, 2002; Gollob & Reichardt, 1991; MacKinnon, 2008; Reichardt, 2011). Figure 2 is an example of a longitudinal extension of the equations presented for the single mediator model with randomized intervention, X, pretest measures of the mediator and outcome, M_1 and Y_1, and posttest measures of the mediator and outcome, M_2 and Y_2, and M_3 and Y_3.

Figure 2.
Three wave Mediation model

Equations 4 – 7 (Valente, 2018) represent the regression equations for the longitudinal mediation model assuming the mediator and outcome variables are continuous variables.

$$M_2 = i_{m2} + a_{m2x}X + s_{m2m1}M_1 + f_{m2y1}Y_1 + e_{m2} \tag{4}$$
$$Y_2 = i_{y2} + c'_{y2x}X + b_{y2m2}M_2 + b_{y2m1}M_1 + s_{y2y1}Y_1 + e_{y2} \tag{5}$$
$$M_3 = i_{m3} + a_{m3x}X + s_{m3m2}M_2 + f_{m3y2}Y_2 + e_{m3} \tag{6}$$
$$Y_3 = i_{y3} + c'_{y3x}X + b_{y3m2}M_2 + b_{y3m3}M_3 + s_{y3y2}Y_2 + e_{y3} \tag{7}$$

Equation 4 represents the effect of the randomized experiment on the time 2 measure of the mediator (a_{m2x}), the effect of the pretest measure of the mediator (M_1) on the time 2 mediator (s_{m2m1}), and the effect of the pretest measure of the outcome (Y_1) on the time 2 mediator (f_{m2y1}). Equation 5 represents the effect of the randomized experiment on the time 2 measure of the outcome (c'_{y2x}), the effect of the time 2 measure of the mediator (M_2) on the time 2 measure of the outcome (b_{y2m2}) the effect of the pretest measure of the mediator (M_1) on the time 2 measure of the outcome (b_{y2m1}), the effect of the pretest measure of the outcome (Y_1) on the time 2 measure of the outcome (s_{y2y1}). Equations 6 – 7 represent the effects from Equations 4 and 5 respectively but from the time 2 measure of the mediator and outcome to time 3 measures of the mediator and outcome. All effects in each equation are adjusted for the other effects in the model and $i_{m2}, i_{y2}, i_{m3},$ and i_{y3} represent intercepts and $e_{m2}, e_{y2}, e_{m3},$ and e_{y3} represent residuals. The total effect of X on Y_2 (c_{y2x}) is obtained by regressing Y_2 on X with no other

predictors in the model and the total effect of X on Y_3 (c_{y3x}) is obtained by regressing Y_3 on X with no other predictors in the model.

The macros presented in this paper are focused on estimating two mediated effects in the three-wave longitudinal mediation model using six different statistical models. *Mediated effect 1* is the mediated effect of the randomized experiment (X) on the outcome at time 2 (Y_2) through its effect on the mediator at time 2 (M_2). Mediated effect 1 represents the two-wave mediated effect which is adjusted for individual differences on the time 1 mediator and outcome (M_1 and Y_1). *Mediated effect 2* is the mediated effect of the randomized experiment (X) on the outcome at time 3 (Y_3) through its effect on the mediator at time 2 (M_2). Mediated effect 2 represents the three-wave *longitudinal* mediated effect because the indirect effect of the X on Y_3 occurs through the mediator at a previous time point (i.e., M_2). Six statistical models are available to estimate the mediated effects in this design. The models are: Analysis of Covariance (ANCOVA), path analysis, difference scores, residualized change scores, Inverse propensity weighting with and without weight truncation, and sequential G-estimation.

STATISTICAL MODELS

ANCOVA and Path Analysis

Mediated effect 1 (X on Y_2 through M_2) is estimated as the product $a_{m2x}b_{y2m2}$ from Equations 4 and 5, respectively, and Mediated effect 2 (X on Y_3 through M_2) is estimated as the product $a_{m2x}b_{y3m2}$ from Equations 4 and 7, respectively (Valente, 2018). Path analysis can estimate the same equations as ANCOVA simultaneously using SAS® PROC CALIS with METHOD=FIML for full-information maximum likelihood. This is the preferred option if there are any missing data (Enders, 2010).

Difference Scores

Equations 8-9 represent the difference scores for the mediator and the outcome, respectively, across time 1 and time 2 and Equation 10 represents the difference score for the outcome across time 2 and time 3 (Valente, 2018).

$$\Delta_{M21} = M_2 - M_1 \tag{8}$$

$$\Delta_{Y21} = Y_2 - Y_1 \tag{9}$$

$$\Delta_{Y32} = Y_3 - Y_2 \tag{10}$$

Equations 11-13 represent regression equations using difference scores for the mediator variable and dependent variable, respectively.

$$\Delta_{M21} = i_{\Delta_{M21}} + a_{\Delta2}X + e_{\Delta_{M21}} \tag{11}$$

$$\Delta_{Y21} = i_{\Delta_{Y21}} + c'_{\Delta2}X + b_{\Delta2}\Delta_{M21} + e_{\Delta_{Y21}} \tag{12}$$

$$\Delta_{Y32} = i_{\Delta_{Y32}} + c'_{\Delta3}X + b_{\Delta3}\Delta_{M21} + e_{\Delta_{Y32}} \tag{13}$$

Mediated effect 1 is estimated as the product of $a_{\Delta2}$ coefficient from Equation 11 and $b_{\Delta2}$ coefficient from Equation 12 ($a_{\Delta2}b_{\Delta2}$) which is the effect of X on change in Y from time 1 to time 2 through its effect on change in M from time 1 to time 2. Mediated effect 2 is estimated as the product of $a_{\Delta2}$ coefficient from Equation 11 and $b_{\Delta3}$ coefficient from Equation 13 ($a_{\Delta2}b_{\Delta3}$) which is the effect of X on change in Y from time 2 to time 3 through its effect on change in M from time 1 to time 2 (Valente, 2018).

Residualized Change Scores

Equation 14 represents residualized change scores calculated for the mediator variable, where R_{M2} indicates change in predicted scores on the mediator variable measured at time 2 subtracted from observed scores on the mediator variable measured at time 2. Equation 15 represents residualized change scores calculated for the dependent variable, where R_{Y2} indicates change in predicted scores on

the dependent variable measured at time 2 subtracted from observed scores on the dependent variable at time 2. Equation 16 represents the residualized change score calculated for the dependent variable where R_{Y32} represents predicted scores on the dependent variable measured at time 3 subtracted from observed scores on the dependent variable measured at time 3 (Valente, 2018).

$$R_{M21} = M_2 - E[M_2|M_1] \tag{14}$$

$$R_{Y21} = Y_2 - E[Y_2|Y_1] \tag{15}$$

$$R_{Y32} = Y_3 - E[Y_3|Y_2] \tag{16}$$

Equations 17-19 represent regression equations using residualized change scores for the mediator variable and outcome variables.

$$R_{M21} = i_{R_{M21}} + a_{R2}X + e_{R_{M21}} \tag{17}$$

$$R_{Y21} = i_{R_{Y21}} + c'_{R2}X + b_{R2}R_{M21} + e_{R_{Y21}} \tag{18}$$

$$R_{Y32} = i_{R_{Y32}} + c'_{R3}X + b_{R2}R_{M21} + e_{R_{Y32}} \tag{19}$$

Mediated effect 1 is estimated by computing the product of a_{R2} coefficient from Equation 17 and b_{R2} coefficient from Equation 18 ($a_{R2}b_{R2}$) which is the effect of X on the change in Y from time 1 to time 2 not predictable from Y_1 through its effect on the change in M from time 1 to time 2 not predictable from M_1. Mediated effect 2 of X on Y_3 through M_2 for the residualized change score model is estimated by computing the product of a_{R2} coefficient from Equation 17 and b_{R2} coefficient from Equation 19 ($a_{R2}b_{R2}$) which is the effect of X on the change in Y from time 2 to time 3 not predictable from Y_2 through its effect on the change in M from time 1 to time 2 not predictable from M_1 (Valente, 2018).

Inverse Propensity Weighting

The potential outcomes framework for causal inference provides nonparametric definitions and assumptions to identify causal effects in randomized experiments and nonrandomized studies (Imbens & Rubin, 2015; Holland, 1986; Rubin, 1974). It is assumed that each unit has a potential value on the outcome and on the mediator for each level of treatment regardless of which treatment group the unit participated in (i.e., potential outcomes). Treatment effects and mediated effects are thus defined as differences between potentially unobservable potential outcomes. This conceptualization of treatment and mediated effects has led to the development of new statistical methods for treatment effect and mediated effect estimation. Inverse propensity weighting (IPW) is a method developed from the potential outcomes framework and is an application of propensity scores (Coffman, 2011; Harder, Stuart, & Anthony, 2010; Hirano, Imbens, & Ridder, 2003; Jo, Stuart, MacKinnon, & Vinokur, 2011; Robins, Hernán, & Brumback, 2000; Stuart, 2010; West et al., 2014). In more recent applications, IPW has been used to account for selection into a given level of treatment over one or more alternatives (Imbens & Rubin, 2015, Chapter 12; Rosenbaum & Rubin, 1983).

When the mediator variable is continuous, it is important to estimate *stabilized* inverse propensity weights (Robins, et al., 2000; Naimi, Moodie, Auger, & Kaufman, 2014). Stabilized weights are created using two propensity scores, a numerator propensity score and a denominator propensity score in the form $f(m_2)/f(m_2|m_1,y_1)$ where f is used to denote the probability (density) of a continuous random variable. The numerator of the weight contains the marginal probability of the mediator, $f(m_2)$, which is the probability of an individual being at their observed value of the mediator with values further away from the mean of the mediator resulting in a *lower* probability of occurring. The denominator of the weight contains the conditional probability of an individual being at their observed level of the mediator at time 2, (e.g., m_2), conditional on their observed levels of pretest measures (e.g., m_1 and y_1). The conditional probability is a predicted probability of being at an observed value of the mediator with values further away from the predicted value resulting in a *lower* probability of occurring.

Stabilized weights are important for continuous treatments because they reduce the sampling variability of the IPW estimator relative to unstabilized weights (Robins, 2000; Robins et al., 2000; Naimi et al.,

2014). Extreme weights (i.e., extremely small or extremely large) increase the sampling variability of the IPW estimator (Cole & Hernán, 2008; Hernán & Robins, 2006). The size of the weights are influenced by the distributional form of the treatment (i.e., treatment is continuous; Naimi et al., 2014) and by the strength of the predictors of the continuous treatment. That is, stronger predictors lead to extreme weights more often than weaker predictors (Goetgeluk, Vansteelandt, & Goetghebeur, 2009; Robins et al., 2000; Vansteelandt, 2009). To remedy extreme stabilized weights, Cole and Hernán (2008) suggested truncation such that extremely large weights (e.g., > 10) are set equal to the 99th, 95th, or 90th percentile of the distribution of weights and extremely small weights (e.g., < 0.01) are set equal to the 1st, 5th, or 10th percentile of the distribution of weights. While truncating stabilized weights reduces the sampling variability of the IPW estimator, doing so increases bias (Cole & Hernán, 2008).

IPW requires two models for each estimated mediated effect. IPW requires a propensity score model and an outcome model for each estimated mediated effect. The propensity score models for Mediated effect 1(X on Y_2 through M_2) is described in Equations 20 – 21 which are used to create the weights described in Equation 24. The outcome model for Mediated effect 1 is described in Equation 26 (Valente, 2018; Valente, MacKinnon, & Mazza, 2018). The propensity score models for Mediated effect 2 (X on Y_3 through M_2) are presented in Equations 22 – 23 which are used to create the weights described in Equation 25. The outcome model for Mediated effect 2 is described in Equation 27 (Valente, 2018).

$$f(M_2|X) = \phi(M_2|X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m_2-\widehat{m}_2)^2}{2\sigma^2}} \tag{20}$$

$$f(M_2|X,M_1,Y_1) = \phi(M_2|X,M_1,Y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m_2-\widehat{m}_2)^2}{2\sigma^2}} \tag{21}$$

$$f(M_3|X) = \phi(M_3|X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m_3-\widehat{m}_3)^2}{2\sigma^2}} \tag{22}$$

$$f(M_3|X,M_2,Y_2) = \phi(M_3|X,M_2,Y_2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m_3-\widehat{m}_3)^2}{2\sigma^2}} \tag{23}$$

Because the mediator at posttest is a continuous variable, the propensity score in the numerator (i.e., $f(M_2|X)$ and denominator are drawn from the normal probability density function (i.e., ϕ) (Robins et al., 2000). m_2 and $\widehat{m}_2$ are the observed and predicted values of the mediator at posttest for a given individual, and σ is the residual standard error from the respective equation. m_3 and $\widehat{m}_3$ are the observed and predicted values of the mediator at time 3 for a given individual, and σ is the residual standard error from the respective equation. The weight used for Mediated effect 1 for each individual is the ratio of the two probabilities from Equations 20 and 21 as in Equation 24 and the weight used for Mediated effect 2 for each individual is the ratio of the two probabilities from Equations 22 and 23 as in Equation 25 (Valente, 2018).

$$w_{m2i} = \frac{f(M_2|X)}{f(M_2|X,M_1,Y_1)}. \tag{24}$$

$$w_{m3i} = \frac{f(M_3|X)}{f(M_3|X,M_2,Y_2)}. \tag{25}$$

Mediated effect 1 is estimated as the difference between the total effect of X on Y_2 (c_{y2x}) and the weighted direct effect from Equation 26 ($c_{y2x(weighted)}$). Mediated effect 2 is estimated as the product of the effect of X on M_2 (a_{m2x}) from Equation 4 and the weighted effect of M_2 on Y_3 ($b_{y3m2(weighted)}$) from Equation 27 (Valente, 2018).

$$E[Y_2|X,M_2] = c'_{y2x(weighted)}X + b_{y2m2(weighted)}M_2 \tag{26}$$

$$E[Y_3|X,M_2,M_3] = c'_{y3x(weighted)}X + b_{y3m2(weighted)}M_2 + s_{y3y2(weighted)}Y_2 \tag{27}$$

Sequential G-estimation

G-estimation relies on removing effects of treatment and mediators from the outcome, Y, *conditional* on covariates and closely resembles regression methods (Vansteelandt, 2009; Vansteelandt & Joffe, 2014). The logic behind G-estimation is that the effect of interest (e.g., direct effect) can be estimated once we remove the effects from the outcome that are not of interest. G-estimation has been applied in to the single mediator model to estimate the direct effect. For example, in the single mediator model, we can think of Y as being a function of a direct effect of X on Y and an indirect effect of X on Y through M. If we remove the indirect effect from Y, then the only remaining effect X could have on Y is the direct effect. The indirect can then be estimated as the difference between the total effect and the G-estimate of the direct effect (Valente, 2018; Valente et al., 2018).

For Mediated effect 1, the adjusted effect of M_2 on Y_2 (b_{y2m2}) from Equation 5 is subtracted from the observed values of Y_2 as in Equation 28. For Mediated effect 2, the adjusted effect of M_2 on Y_3 (b_{y3m2}) from Equation 7 is subtracted from observed values of Y_3 as in Equation 29 (Valente, 2018).

$$Y_{2diff} = Y_2 - b_{y2m2}M_2 \tag{28}$$

$$Y_{3diff} = Y_3 - b_{y3m2}M_2 \tag{29}$$

For Mediated effect 1, the residualized outcome Y_{2diff} from Equation 28 is regressed on the treatment X to estimate an adjusted direct effect (Valente et al., 2018). For Mediated effect 2, the residualized outcome Y_{3diff} from Equation 29 is regressed on the treatment X to estimate an adjusted direct effect.

$$E[Y_{2diff}|X] = c'_{y2x(adjusted)}X \tag{30}$$

$$E[Y_{3diff}|X] = c'_{y3x(adjusted)}X \tag{31}$$

Assuming no interactions, Mediated effect 1 is estimated as $c_{y2x} - c'_{y2x(adjusted)}$ and Mediated effect 2 is estimated as $c_{y3x} - c'_{y3x(adjusted)}$ (Valente, 2018).

Effect Size

An effect size estimate is computed for each mediated effect for each model by dividing the respective mediated effect by the standard deviation of the outcome (Miočević, O'Rourke, MacKinnon, & Brown, 2017). For ANCOVA, Path analysis, IPW, and sequential G-estimation, the effect size estimate for Mediated effect 1 is (Mediated effect 1/standard deviation of Y_2) and the effect size estimate for Mediated effect 2 is (Mediated effect 2/standard deviation of Y_3). For the difference score model, the effect size estimate for Mediated effect 1 is (Mediated effect 1/standard deviation of Δ_{Y21}) and the effect size estimate for Mediated effect 2 is (Mediated effect 2/standard deviation of Δ_{Y32}). For the residualized change score model, the effect size estimate for Mediated effect 1 is (Mediated effect 1/standard deviation of R_{Y21}) and the effect size estimate for Mediated effect 2 is (Mediated effect 2/standard deviation of R_{Y32}).

The macros in this paper are extensions of three existing SAS[®] macros %TWOWAVEMED, %TWOWAVEMONTECARLO, and %TWOWAVEPOSTPOWER for the estimation of mediated effects in a pretest-posttest control group design (Valente & MacKinnon, 2017a; 2017b). The macros presented in this paper are the %THREEWAVEMED and %THREEWAVEMONTECARLO. Models are estimated using SAS® PROC REG, PROC CALIS, PROC GENMOD, and DATA steps when necessary.

CAUSAL LONGITUDINAL MEDIATION MACROS

THE THREEWAVEMED MACRO

%***threewavemed***(data=,x=,m1=,y1=,m2=,y2=,m3=,y3=,cov1=,cov2=,cov3=,nobs=,boot=true,nboot=**1000**,alpha=**.05**);

To use the threewavemed macro, users enter the name of a SAS dataset of their choice (*data* =), specify the name of the randomized experimental variable (x=,) specify the names of the continuous mediating and outcome variables at each of the three waves (m1=,y1=,m2=,y2=,m3=,y3=,) and specify the names of up to three baseline covariates in addition to the wave 1 mediating and outcome variables (cov1=,cov2=,cov3=,). The dataset that is specified by the user is required to be in wide format. That is, each variable at each time point should be a separate column in the dataset. Next, users' specify the number of observations in the dataset (nobs=,) whether they want single sample significance results or percentile bootstrap significance results (boot =), number of bootstrap samples (nboot =), and alpha level (alpha =).

Single sample results refer to normal theory significance test and asymmetric confidence intervals (i.e., PRODCLIN; MacKinnon, Fritz, Williams, & Lockwood, 2007). The preferred single sample method is the asymmetric confidence intervals as they provide a more accurate test of the mediated effect because the product of two regression coefficients is not normally distributed for finite sample sizes and a range of effect sizes (MacKinnon, Lockwood, & Williams, 2004). When single sample results are requested by the user, the program makes use of the PRODCLIN program (MacKinnon et al., 2007) (instructions for download are included with the SAS code for these macros). If single sample tests are requested (i.e., boot=false), users will get estimates of Mediated effect 1 and Mediated effect 2 for each of the four following models: ANCOVA, Path analysis, difference scores, and residualized change scores. The potential outcomes methods, IPW and sequential G-estimation, and confidence intervals for the effect sizes estimates are only available when bootstrap significance tests are specified (i.e., boot=true).

Tests of the Mediated effect with distribution of a product confidence limits Mediated Effect 1 for α=.05

Estimator	Estimate	Effect size	z-value	p-value	Lower Limit	Upper Limit
ANCOVA	0.20200	0.20060	5.325	.0000001	0.13219	0.28087
Path	0.20200	0.20060	5.350	.0000000	0.13249	0.28048
Diff	0.08629	0.07228	2.671	.0037844	0.02749	0.15545
Res	0.17790	0.19047	4.985	.0000003	0.11281	0.25308

Tests of the Mediated effect with distribution of a product confidence limits Mediated Effect 2 for α=.05

Estimator	Estimate	Effect size	z-value	p-value	Lower Limit	Upper Limit
ANCOVA	0.082000	0.074323	2.494	0.00631	0.02049	0.15071
Path	0.082000	0.074323	2.507	0.00609	0.02080	0.15035
Diff	-0.060007	-0.056024	-2.093	0.98183	-0.12093	-0.00690
Res	0.054870	0.056917	1.896	0.02899	0.00026	0.11529

Output 1. Tests of Mediated effect 1 and Mediated effect 2 using distribution of a product confidence limits with estimates of effect size. ANCOVA = Analysis of Covariance, Path = Path analysis, Diff = Difference scores, Res = Residualized change scores.

If users have missing data on the posttest waves of mediating and outcome variables, they have three options. First, users can create a listwise deleted dataset assuming the data is missing completely at random. If users choose to use a listwise deleted dataset, then (nobs =) would be specified to equal the size of the listwise deleted dataset. Second, users can impute the missing values of the posttest mediating and outcome variables based on pretest measures of the mediating and outcome variables and any other baseline covariates assuming the data are missing at random. If users choose to use an imputed dataset, then (nobs =) would be specified to equal the size of the imputed dataset. Third, assuming the data are missing at random, users can choose (boot=false) option and report the path analysis results because METHOD=FIML (i.e., full-information maximum likelihood) is specified when using SAS® PROC CALIS (Enders, 2010). The first two options will work with both the single sample procedures and the bootstrap procedures. The third option will only work with the single sample path analysis procedure.

If users specify baseline covariates (cov1 = , cov2 = , cov3 =), it is assumed that the baseline covariates are correlated with each other and the baseline measures of the mediator and the outcome (i.e., M_1 and Y_1) and uncorrelated with X. Further, it is assumed the baseline covariates only affect the wave 2 mediator and outcome variables (i.e., M_2 and Y_2) and not the wave 3 mediator and outcome variables (i.e., M_3 and Y_3). Therefore, baseline covariates are only included in the statistical models for relations between time 1 and time 2.

Estimate of the Mediated effect with percentile bootstrap confidence limits
Mediated Effect 1
500 bootstrap samples and α=.05

Estimator	Estimate	Lower Limit	Upper Limit	Effect size	Lower Limit	Upper Limit
Ancova	0.20152	0.12351	0.29187	0.20060	0.12318	0.28127
Path	0.20152	0.12351	0.29187	0.20060	0.12318	0.28127
Diff	0.08629	0.01219	0.16739	0.07228	0.01002	0.14059
Res	0.17790	0.10212	0.26190	0.19047	0.10975	0.27555
IPW	-0.07550	-0.39826	0.30210	-0.07515	-0.40471	0.30491
IPW-99	0.11917	-0.27974	0.37992	0.11862	-0.27633	0.37125
IPW-95	0.20148	0.07171	0.33867	0.20056	0.07531	0.32651
IPW-90	0.20934	0.11341	0.32197	0.20838	0.11572	0.31715
Seq-g	0.18745	0.10860	0.28351	0.18659	0.10910	0.27222

Output 2. Tests of Mediated effect 1 and Mediated effect 2 using percentile bootstrapping with estimates and percentile bootstrap confidence intervals for effect size. ANCOVA = Analysis of Covariance, Path = Path analysis, Diff = Difference scores, Res = Residualized change scores, IPW = Inverse Propensity Weighting, IPW-99 = IPW with weights truncated at the 1st and 99th percentiles, IPW-95 = IPW with weights truncated at the 5th and 95th percentiles, IPW-90 = IPW with weights truncated at the 10th and 90th percentiles.

|	Estimate of the Mediated effect with percentile bootstrap confidence limits Mediated Effect 2 500 bootstrap samples and α=.05					
Estimator	**Estimate**	**Lower Limit**	**Upper Limit**	**Effect size**	**Lower Limit**	**Upper Limit**
Ancova	0.082047	0.02140	0.14878	0.074323	0.01902	0.13457
Path	0.082047	0.02140	0.14878	0.074323	0.01902	0.13457
Diff	-0.060007	-0.12311	-0.00898	-0.056024	-0.11280	-0.00783
Res	0.054870	0.00949	0.10932	0.056917	0.00991	0.11668
IPW	0.078026	0.00237	0.16439	0.070681	0.00221	0.14801
IPW-99	0.077975	0.00179	0.16439	0.070635	0.00151	0.14881
IPW-95	0.082341	0.01828	0.15660	0.074590	0.01588	0.14394
IPW-90	0.098766	0.04067	0.17529	0.089469	0.03686	0.15831
Seq-g	0.076319	0.02081	0.14286	0.069134	0.01850	0.13336

Output 2. Tests of Mediated effect 1 and Mediated effect 2 using percentile bootstrapping with estimates and percentile bootstrap confidence intervals for effect size. ANCOVA = Analysis of Covariance, Path = Path analysis, Diff = Difference scores, Res = Residualized change scores, IPW = Inverse Propensity Weighting, IPW-99 = IPW with weights truncated at the 1st and 99th percentiles, IPW-95 = IPW with weights truncated at the 5th and 95th percentiles, IPW-90 = IPW with weights truncated at the 10th and 90th percentiles. *Continued.*

THE THREEWAVEMONTECARLO MACRO

%***threewavemontecarlo***(model=,
by1m1=,am2x=,by2m2=,cy2x=,sm2m1=,sy2y1=,fm2y1=,by2m1=,
am3x=,by3m2=,by3m3=,cy3x=,sm3m2=,sy3y2=,fm3y2=,
varm1=,vary1=,varm2=,vary2=,varm3=,vary3=,
begin=,end=,incr=,nsim=,boot=,nboot=,alpha=);

To use the threewavemontecarlo macro, users specify which of the six statistical models they want to estimate empirical power for (model =), estimates of the path estimates and error variances from Figure 2 (by1m1=,am2x=,by2m2=,cy2x=,sm2m1=,sy2y1=,fm2y1=,by2m1=,am3x=,by3m2=,by3m3=, cy3x=,sm3m2=,sy3y2=,fm3y2=,varm1=,vary1=,varm2=,vary2=,varm3=,vary3=,), the sample size the user wants the power analysis to begin with (begin =), the sample size the user wants the power analysis to end at (end =), the sample size increment (incr =)for which to increase the starting sample size by until it reaches the end sample size. Users also need to specify the number of simulations per sample size (nsim =), whether the user wants empirical power estimates based on single sample test (boot = false) or empirical power based on a percentile bootstrap procedure (boot = true), the number of bootstrap replications per sample size per simulation (nboot =), and the desired alpha level (alpha =). The user can choose to estimate empirical power for a single sample size by specifying the beginning, end, and increment samples sizes to be equal (e.g., begin=50, end=50, incr=50). Both single sample and percentile bootstrap empirical power can be estimated for the following models: ANCOVA, Path analysis, difference scores, residualized change scores. Only percentile bootstrap empirical power is available for the potential outcomes models, IPW and sequential G-estimation. That is, (boot = true) must be specified to obtain empirical power estimates for the potential outcomes models.
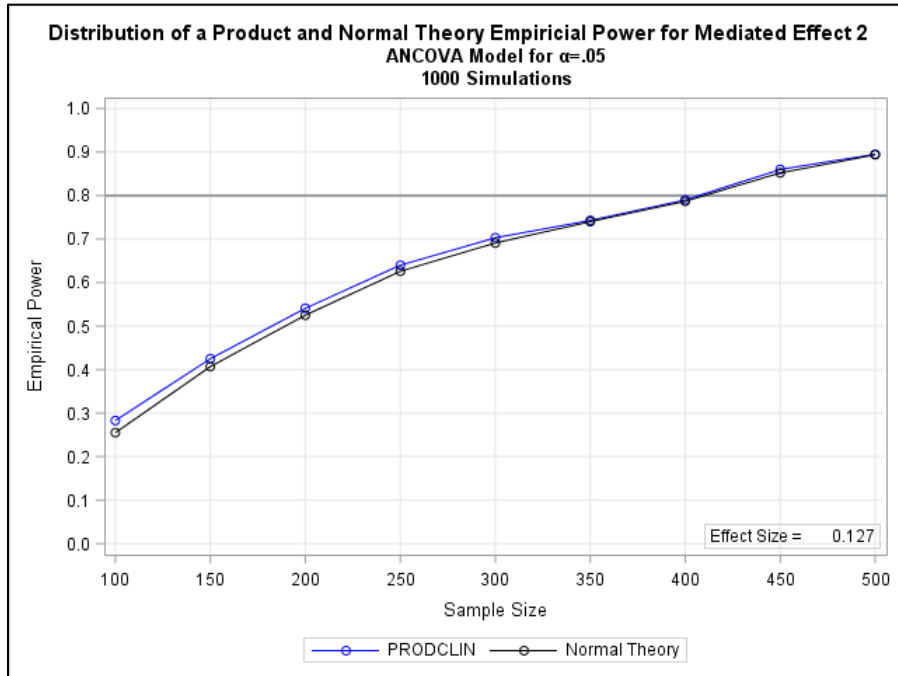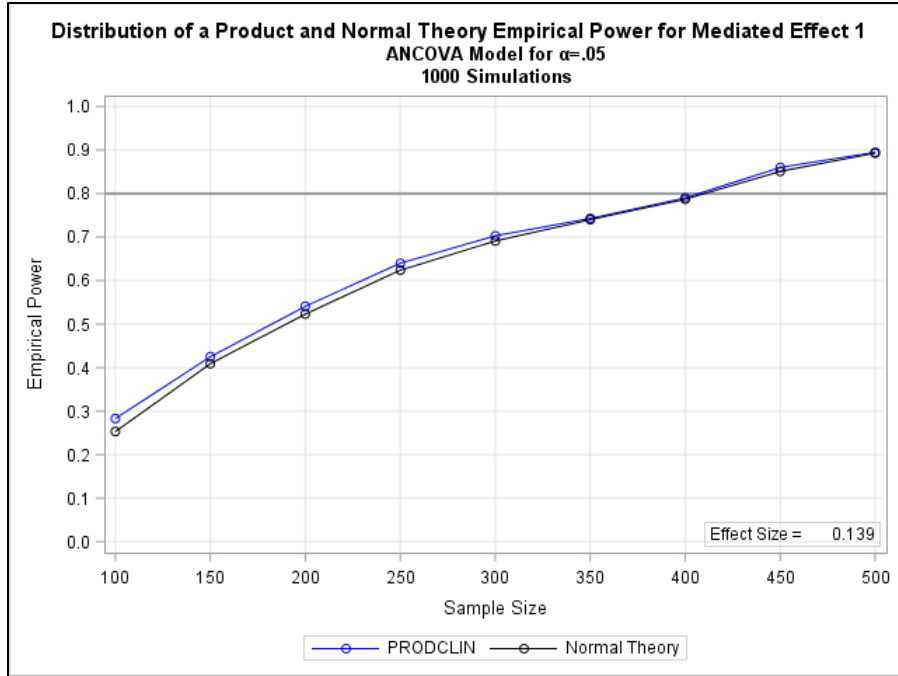
Output 3. Prospective power analysis using normal theory and distribution of a product methods to test Mediated effect 1 and Mediated effect 2. Empirical power plotted by sample size with reference line at 0.80 power.
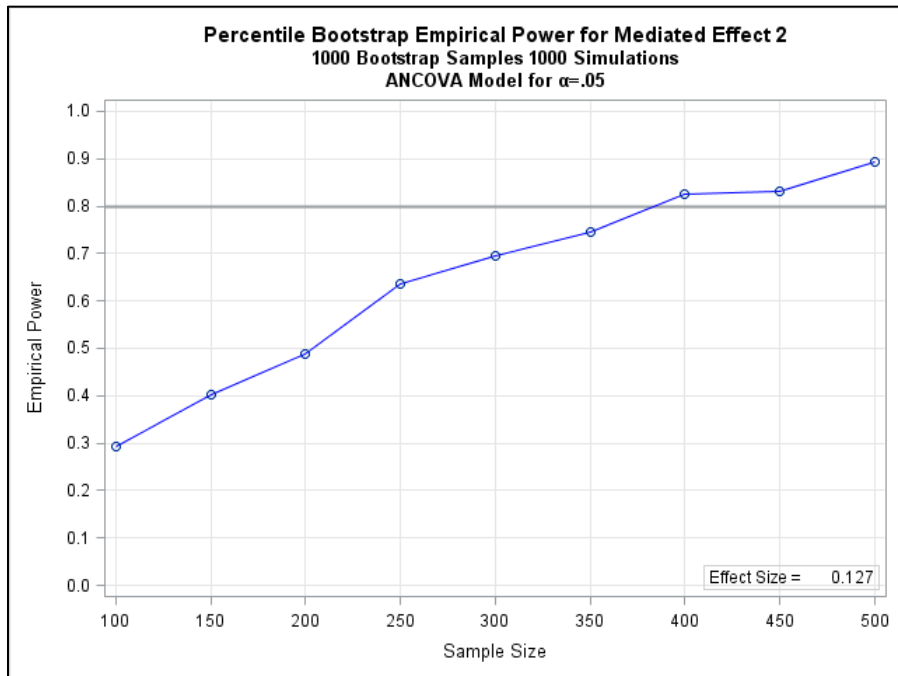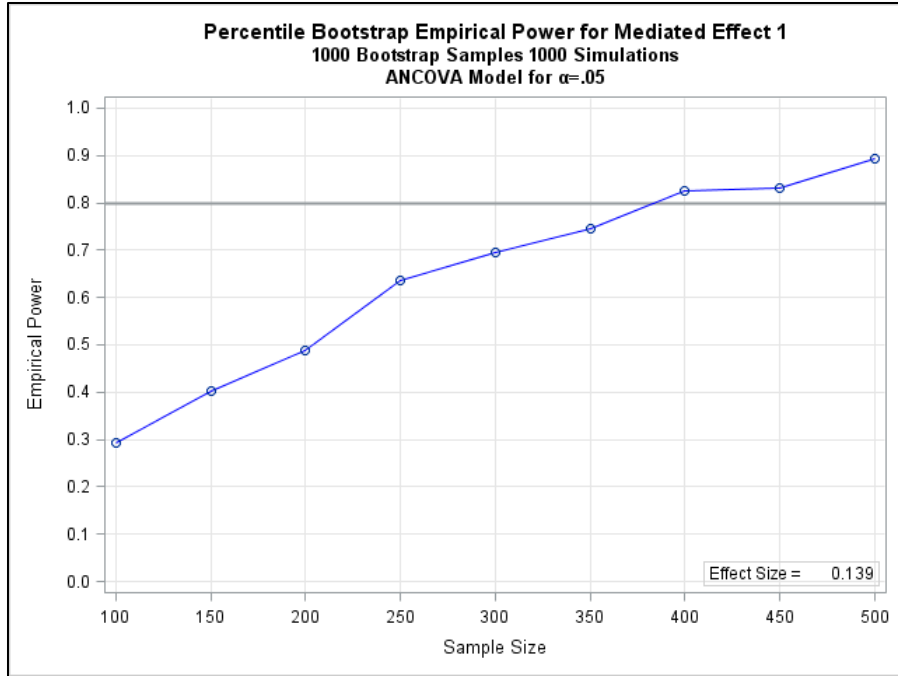
Output 4. Prospective power analysis using percentile bootstrap confidence intervals to test Mediated effect 1 and Mediated effect 2. Empirical power plotted by sample size with reference line at 0.80 power.

The empirical power plots will appear smoother and continuously increasing as a function of sample size if either the simulation replications or the bootstrap samples are increased. Some researchers recommend using 10,000 simulation replications for Monte Carlo power estimates (Thoemmes, MacKinnon, & Resier 2010).

If users are interested in conducting a post-hoc power analysis, they need to specify the path estimates and error variances from an observed dataset and specify the beginning, end, and increment samples sizes to be equal (e.g., begin=50, end=50, incr=50).

Distribution of a Product and Normal Theory Empirical Power for Mediated Effect 1
Sample Size = 500 Simulations = 1000
ANCOVA Model for α = .05

Mediated Effect 1 E.S.	PRODCLIN Power	Normal Theory Power
0.139	0.892	0.886

Distribution of a Product and Normal Theory Empiricial Power for Mediated Effect 2
Sample Size = 500 Simulations = 1000
ANCOVA Model for α = .05

Mediated Effect 2 E.S.	PRODCLIN Power	Normal Theory Power
0.127	0.892	0.885

Output 5. Retrospective empirical power analysis using normal theory and distribution of a product confidence intervals for the ANCOVA model estimate of Mediated effect 1 and Mediated effect 2.

Percentile Bootstrap Empirical Power for Mediated Effect 1
Sample Size = 500 Simulations = 1000 Bootstrap Samples = 1000
ANCOVA Model for α = .05

Mediated Effect 1 E.S.	Percentile Boostrap Power
0.139	0.897

Percentile Bootstrap Empirical Power for Mediated Effect 2
Sample Size = 500 Simulations = 1000 Bootstrap Samples = 1000
ANCOVA Model for α = .05

Mediated Effect 2 E.S.	Percentile Boostrap Power
0.127	0.897

Output 6. Retrospective empirical power analysis using percentile bootstrap confidence intervals for the ANCOVA model estimate of Mediated effect 1 and Mediated effect 2.

Output 5 and Output 6 correspond to the retrospective power given the ANCOVA estimate of the mediated effect. Retrospective power is typically used when a significance test fails to reject the null hypothesis and a researcher wants to know what the power was given the sample estimated effect size to find the effect. A drawback of this approach is the assumption that the sample estimate of effect size is treated as the actual population value of the effect size which is likely not true and can lead to the "power approach paradox" (Hoenig & Heisey, 2001).

CONCLUSION

Researchers can use the two macros presented to estimate two mediated effects in a three-wave longitudinal mediation model, conduct sample size planning for new research studies involving the mediated effects in the three-wave mediation model and determine retrospective power for observed mediated effects from previous research studies. These macros provide users with tools for all stages of design and analysis from prospective power analyses (THREEWAVEMONTECARLO), to estimating mediated effects in observed data (THREEWAVEMED), and probing retrospective power if prospective power analyses were not originally conducted (THREEWAVEMONTECARLO). Users can obtain the macro syntax by emailing the first author or visiting the first author's ResearchGate profile which can be found in the contact information at the end of this paper. The simplest way to use these macros is to copy and paste the macro syntax into an existing SAS® syntax file, run the macro code, and then call the respective macro (e.g., %threewavemed(...)) and specify the user inputs. Overall, these macros will provide users with powerful analytical tools for this popular experimental design while harnessing the reliable and efficient use of SAS® PROCs.

REFERENCES

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology, 98*(4), 550-558.

Coffman, D. L. (2011). Estimating causal effects in mediation analysis using propensity scores. *Structural Equation Modeling, 18*(3), 357-369.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*(4), 558.

Cole, S. R., & Hernán, M A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology, 168*(6), 656-665.

Collins, L. M., & Graham, J. W. (2002). The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. *Drug and Alcohol Dependence, 68* (1), 85-96.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY:Guilford Press.

Goetgeluk, S., Vansteelandt, S., & Goetghebeur, E. (2009). Estimation of controlled direct effects. *Journal of the Royal Statistical Society: Series B, 70*(5) 1049-1066.

Gollob, H. F., & Reichardt, C. S. (1991). Interpreting and estimating indirect effects assuming time lags really matter. *Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions.* American Psychological Association, Washington, DC.

Harder, V .S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234-249.

Hernán, M. A., & Robins, J. M. (2006). Estimating causal effects from epidemiological data, *Journal of Epidemiology and Community Health*, *60*(7), 578–586.

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica, 71*(4), 1161-1189.

Hoenig, J. M., & Heisey, D. M. (2012). The abuse of power. *The American Statistician*, *55*(1), 19-24.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945-960.

Holland, P. W. (1988). Causal inference, path analysis and recursive structural equations models (with discussion). In C. Clogg (Ed.), *Sociological Methodology* (pp. 449 – 484). Washington, DC: American Sociological Association.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309-334.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational Studies. *American Political Science Review*, *105*(4), 765–789.

Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference in statistics, social, and biomedical sciences*. New York, NY: Cambridge University Press.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, *69*(2), 307.

Jo, B., Stuart, E. A., MacKinnon, D. P., & Vinokur, A. D. (2011). The use of propensity scores in mediation analysis. *Multivariate Behavioral Research, 46*(3), 425-452.

Larimer, M. E., Lee, C. M., Kilmer, J. R., Fabiano, P. M., Stark, C. B., Geisner, I. M., … Neighbors, C. (2007). Personalized feedback for college drinking prevention: A randomized control trial. *Journal of Consulting and Clinical Psychology, 75*(2), 285-293.

Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In P.F. Lazarsfeld & M. Rosenburg (Eds.), *The language of social research: A reader in the methodology of social research* (pp. 115 – 125). Glencoe, IL: Free Press

MacKinnon, D. P. (1994). Analysis of mediating variables in prevention and intervention research. In A. Cazares & L. A. Beatty (Eds.), *Scientific methods for prevention/intervention research* (NIDA Research Monograph Series 139, DHHS Pub 94-3631, pp. 127-153). Washington DC: U. S. Department of Health and Human Services.

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

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, *17*(2), 144-158

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods*, *39*(3), 384-389.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate behavioral research*, *39*(1), 99-128.

MacKinnon, D. P., Goldberg, L., Clarke, G. N., Elliot, D. L., Cheong, J., Lapin, A., Moe, E. L., & Krull, J. L. (2001). Mediating mechanisms in a program to reduce intentions to use anabolic steroids and improve exercise self-efficacy and dietary behavior. *Prevention Science, 2*(1), 15-28.

MacKinnon, D. P., & Pirlott, A. G. (2015). Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. *Personality and Social Psychology Review, 19*(1), 30-43.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23.

Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research*, *46*(5), 816-841.

Miočević, M., O'Rourke, H. P., MacKinnon, D. P., & Brown, C. H. (2017). Statistical properties of five effect size measures for mediation models. *Behavior Research Methods.* Advance Online Publication.

Naimi, A. I., Moodie, E. E. M., Auger, N., & Kaufman, J. S. (2014). Constructing inverse probability weights for continuous exposures. *Epidemiology*, *25*(2), 292–299.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles (with discussion). Section 9 (D. M. Dabrowska & T. P. Speed, Trans.). *Statistical Science*, 5, 465-480.

Pearl, J. (2001). Direct and indirect effects. In J. Breese & D. Koller (Eds.), *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence* (pp. 411-420). San Francisco, CA: Morgan Kaufmann.

Pearl, J. (2009). *Causality*. New York, NY: Cambridge university press.

Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods*, *19*(4), 459-481.

Reichardt, C. S. (2011). Commentary: Are three waves of data sufficient for assessing mediation? *Multivariate Behavioral Research*, *46*(5), 842–851.

Robins, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In *Statistical models in epidemiology, the environment, and clinical trials* (pp. 95-133). New York, NY: Springer.

Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology, 3*(2), 143-155.

Robins, J. M., Hernán, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology, 11*(5), 550-560.

Rosenbaum, B. Y. P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *7*(1), 41–55.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688-701.

Sobel, M. E. (1990). Effect analysis and causation in linear structural equation models. *Psychometrika, 55*(3), 495-515.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*(1), 1-21.

Thoemmes, F., MacKinnon, D. P., & Reiser, M. R. (2010). Power analysis for complex mediational designs using Monte Carlo methods. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(3), 510-534.

Valente, M. J. (2018). *Comparison of Methods for Estimating Longitudinal Indirect Effects.* (Unpublished doctoral dissertation). Arizona State University, Tempe AZ.

Valente, M. J. & MacKinnon, D. P (2017a). Comparing models of change to estimate the mediated effect in the pretest-posttest control group design. *Structural Equation Modeling: A Multidisciplinary Journal*.

Valente, M. J., & MacKinnon, D. P. (2017b). SAS® Macros for Computing the Mediated Effect in the Pretest Posttest Control Group Design. *In Proceedings of the SAS Global Forum 2017 Conference, Orlando FL: SAS Institute Inc.*

Valente, M. J., MacKinnon, D. P., & Mazza, G. (2018). Performance of inverse propensity weighting and sequential G-estimation for assessing mediation in the pretest-posttest control group design. *Manuscript in preparation.*

VanderWeele, T., & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface*, *2*(4), 457-468.

Vansteelandt, S. (2009). Estimating direct effects in cohort and case-control studies. *Epidemiology, 20*(6), 851-860.

Vansteelandt, S., & Joffe, M. (2014). Structural nested models and G-estimation: The partially realized promise. *Statistical Science*, *29*(4), 707–731.

Wan, E. W., Rucker, D. D., Tormala, Z. L., & Clarkson, J. J. (2010). The effect of regulatory depletion on attitude certainty. *Journal of Marketing Research*, *47*(3), 531-541.

West, S. G., & Aiken, L. S. (1997). Towards understanding individual effects in multiple component prevention programs: Design and analysis strategies. In K. Bryant, M. Windle & S. G. West (Eds). *Recent advances in prevention methodology: Alcohol and substance abuse research* (pp. 167-209). Washington, DC: American Psychological Association.

West, S. G., Cham, H., Thoemmes, F., Renneberg, B., Schulze, J., & Weiler, M. (2014). Propensity scores as a basis for equation groups: Basis principles and application in clinical treatment outcome research. *Journal of Consulting and Clinical Psychology, 82*(5), 906-919.

Witkiewitz, K., & Bowen, S. (2010). Depression, craving, and substance use following a randomized trial of mindfulness-based relapse prevention. *Journal of Clinical and Consulting Psychology, 78*(3), 362-374.

ACKNOWLEDGMENTS

RECOMMENDED READING

- *Base SAS$^®$ Procedures Guide*
- *SAS$^®$ For Dummies$^®$*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Matthew J. Valente
Organization: Arizona State University
Email: m.valente@asu.edu
Website: https://www.researchgate.net/profile/Matthew_Valente

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.