Abstract

Scientific Abstract: Models where some part of a mediation is moderated (conditional process models) are commonly used in psychology research, allowing for better understanding of when the process by which a focal predictor affects an outcome through a mediator depends on moderating variables. Methodological developments in conditional process analysis have focused on between-subject designs. However, two-instance repeated-measures designs, where each subject is measured twice: once in each of two instances, are also very common. Research on how to statistically test mediation, moderation, and conditional process models in these designs has lagged behind. Judd, Kenny, and McClelland (2001) introduced a piecewise method for testing for mediation, that Montoya and Hayes (2017) then translated to a path-analytic approach, quantifying the indirect effect. Moderation analysis in these designs has been described by Judd, McClelland, and Smith (1996), Judd et al. (2001), and Montoya (2018). The generalization to conditional process analysis remains incomplete. I propose a general conditional process model for two-instance repeated-measures designs with one moderator and one mediator. Simplifications of this general model correspond to more commonly used moderated mediation models, such as first-stage and second-stage conditional process analysis. An applied example shows both how to conduct the analysis using MEMORE, a free and easy-to-use macro for SPSS and SAS, and how to interpret the results of such an analysis. Alternative methods for evaluating moderated mediation in two-instance repeated-measures designs using multilevel approaches are also discussed.

Translational Abstract: Mediation analysis is used to understand processes that underlie causal effects (questions of why or how), and moderation analysis is used to understand contingencies (question of when or for whom). These two methods can be combined in conditional process analysis to evaluate whether the process by which one variable affects another depends on a moderator variable. While these types of questions are paramount to developing theory in psychological science, the statistical methods for conducting these analyses have been limited to designs where individuals are either observed on or randomly assigned to the primary causal variable of interest (between-subject designs). By contrast repeated-measures designs, where participants respond to multiple levels of the independent variable (e.g., responding to multiple types of stimuli, or before and after a treatment), are also common within psychology research. Recent research has developed methods for mediation and moderation in these designs, but not their combination: conditional process analysis. In this paper, I propose a method for estimating a variety of conditional process models, provide an easy-to-use statistical tool for SPSS and SAS to conduct these analyses, and demonstrate the proposed method using a real example. The example looks at why psychological treatments for pain may be less effective for individuals with high inflammation. This new method allows researchers who use repeated-measures designs to pursue theoretically interesting questions related to how processes depend on other factors.

Keywords: Conditional Process Analysis, Moderated Mediation, Repeated-Measures Designs; Linear Regression; Multilevel Models; Structural Equation Models

Conditional Process Analysis for Two-Instance Repeated-Measures Designs

Introduction

Mediation, moderation, and conditional process analysis address common questions in psychological research. Mediation is the hypothesis that a focal predictor indirectly affects an outcome through a mediator variable. Statistical mediation analysis estimates the indirect effect, which is of primary interest to address questions of why or how some effect occurs. Moderation, by contrast, helps researchers evaluate if the relationship between a focal predictor and a dependent variable depends on a moderator. Moderation analysis addresses questions of when or for whom certain effects occur. Conditional process analysis combines mediation and moderation, examining if the effect of a focal predictor (X) on a dependent variable (Y) through a mediator (M) depends on a moderator (W). This analysis helps researchers evaluate whether an indirect effect is consistent across the range of the moderator or if the indirect effect depends on the value of (i.e., is conditional on) the moderator. In this article the term conditional is used to mean evaluated at a specific value of another variable. This is in contrast to "main effects" or "partial effects" which are evaluated by averaging over or controlling for other variables. Thus conditional process analysis gets it's name by its focus on estimating conditional indirect effects and testing whether these effects significantly differ from each other (i.e., Is the mediation moderated?).

Recent developments in conditional process analysis have been primarily focused on analysis with data collected using between-subjects designs (Hayes, 2018b, 2022; Valente, Rijnhart, & Gonzalez, 2023), where each participant has one observation on X, M, Y, and W. However, another very common design in psychology is the two-instance repeated-measures design. In this design each subject is measured on the outcomes twice, once in each of two instances, where the repeated-measures factor is the focal predictor of interest (X) which varies across the two instances. Research on how to statistically test mediation, moderation, and conditional process models in these designs is available but limited: Judd et al. (2001) introduced a piecewise method for testing for mediation, and

Montoya and Hayes (2017) translated it to a path-analytic approach, allowing for a quantification of the indirect effect. Moderation for two-instance repeated-measures designs had been described by Judd et al. (1996), Judd et al. (2001), and Montoya (2019). However, the generalization to conditional process analysis, or moderated mediation, remains incomplete. The purpose of this paper is to propose a general method for testing moderated mediation in two-instance repeated-measures designs, accompanied by an easy-to-use software, MEMORE, for implementing the proposed approach.

Conditional Process Analysis in Between-Subject Designs

In this manuscript, the term *between-subjects designs* is used to describe two common designs in psychological research: (1) between-subjects experimental designs and (2) cross-sectional designs. In a between-subjects experimental design, participants are randomly assigned to a condition (X) and then measured on the mediator (M) and the outcome variable (Y), and the moderator (W) is either randomly assigned or measured in a way that it would not be affected by X. In cross-sectional designs there is no random assignment, but all variables are measured once including the focal predictor of interest (X), the mediator (M), the moderator (W), and the outcome variable (Y). Both between-subjects experimental designs and cross-sectional designs result in similar data structures, which is why they are discussed together in this paper under the common term between-subject designs.

Early approaches to moderated mediation focused on a general model where the moderator is allowed to moderate all three paths in the mediation model: the relationship between X and M, the relationship between M and Y, and the direct effect of X on Y (Fairchild & MacKinnon, 2009; Muller, Judd, & Yzerbyt, 2005). In this framework, regression models are estimated such that each effect is conditional on a common moderator. In the notation used below, if a relationship is conditional on a variable, that variable is included in parentheses, to indicate that the relationship is a function of that variable. By contrast, variables which are controlled for are indicated with a comma in the subscript. For example, $\theta_{X\to Y,M}(W_i)$ would be read as the effect of X on Y controlling for M and conditional on W.

$$
\begin{aligned}
M_i &= a_0 + \theta_{X\to M}(W_i)X_i + a_2W_i + e_{M_i} && (1)\\
Y_i &= c'_0 + \theta_{X\to Y,M}(W_i)X_i + \theta_{M\to Y,X}(W_i)M_i + b_2W_i + e_{Y_i} && (2)
\end{aligned}
$$

where

$$
\begin{aligned}
\theta_{X\to M}(W_i) &= a_1 + a_3W_i\\
\theta_{X\to Y,M}(W_i) &= c'_1 + c'_3W_i\\
\theta_{M\to Y,X}(W_i) &= b_1 + b_3W_i
\end{aligned}
$$

In this model the effect of X on M is a linear function of W, $\theta_{X\to M}(W_i)$, the effect of X on Y controlling for M is a linear function of W, $\theta_{X\to Y,M}(W_i)$, and the effect of M on Y controlling for X is a linear function of W, $\theta_{M\to Y,X}(W_i)$. All three effects are conditional on W: no single coefficient represents the relationship between any two variables. If a_3, c'_3, and b_3 were all zero, then this model would simplify to a mediation model where W is controlled for in all of the equations but does not moderate any of the paths. The indirect effect is the product of the effect of X on M and the effect of M on Y controlling for X. In this general conditional process model, both of these effects are conditional on W, and so the indirect effect will also be conditional on W. I denote the conditional indirect effect as $\theta_{X\to M\to Y}(W)$.

$$
\begin{aligned}
\theta_{X\to M\to Y}(W) &= \theta_{X\to M}(W)\theta_{M\to Y,X}(W)\\
&= (a_1 + a_3W)(b_1 + b_3W)\\
&= a_1b_1 + (a_1b_3 + a_3b_1)W + a_3b_3W^2
\end{aligned}
$$

Though each component of the indirect effect is a linear function of W, the indirect effect is a quadratic function of W. Fairchild and MacKinnon (2009) introduced a test of moderation for this model, if the moderator is dichotomous. This is achieved by taking the difference between the two conditional indirect effects at the coded values of the moderator (w_1 and w_2):

$$\begin{aligned}
\theta_{X\to M\to Y}(W=w_1) &= a_1b_1 + (a_1b_3 + a_3b_1)w_1 + a_3b_3(w_1)^2 \\
\theta_{X\to M\to Y}(W=w_2) &= a_1b_1 + (a_1b_3 + a_3b_1)w_2 + a_3b_3(w_2)^2 \\
\theta_{X\to M\to Y}(W=w_1) - \theta_{X\to M\to Y}(W=w_2) &= (a_1b_3 + a_3b_1)(w_1 - w_2) + (a_3b_3)(w_1^2 - w_2^2)
\end{aligned} \quad (3)$$

Fairchild and MacKinnon (2009) show that a test on this value in Equation 3 is a test of moderated mediation if the moderator is dichotomous. When the moderator is continuous the process of making inference for such a general model is not clear. Some have suggested that if either $a_1b_3 + a_3b_1$ or a_3b_3 are significantly different from zero, moderated mediation is supported (Fairchild & MacKinnon, 2009; Muller et al., 2005). However this relies on a fairly piecewise logic that is not completely satisfactory (Hayes, 2015).

In this general model, all paths are moderated, which may mean it includes potentially unnecessary parameters. This can reduce power to detect moderated mediation effects and make interpretation more complex (Fossum, 2023). Additionally, inference about moderated mediation with these models can be quite difficult because the indirect effect is a nonlinear function of the moderator. The next section examines two simpler models where the conditional indirect effect is a linear function of the moderator, and there exists a test moderated mediation when the moderator is continuous or dichotomous.

First-Stage Conditional Process Analysis. A first stage conditional process model allows the relationship between X and M to be conditional on W, the relationship between M and Y is not conditional on W, and the direct effect could be conditional on W or not. For this type of model the equation for the mediator matches that of the general model (Equation 1). The conditional effect of X on M is $\theta_{X\to M}(W) = a_1 + a_3W$. There are three options for the model of the outcome: 1) the direct effect is moderated by W, 2) the direct effect is not moderated, but W is included as a covariate, and 3) W is not included in the model for Y. These three options are expressed respectively as

$$\begin{aligned} Y_i &= c_0' + (c_1' + c_3' W_i) X_i + b_1 M_i + b_2 W_i + e_{Y_i} \\ Y_i &= c_0' + c_1' X_i + b_1 M_i + b_2 W_i + e_{Y_i} \\ Y_i &= c_0' + c_1' X_i + b_1 M_i + e_{Y_i} \end{aligned}$$

In any of the above cases, the effect of M on Y controlling for X is not conditional on W, it is simply b_1. The indirect effect is the product of the effect of X on M and the effect of M on Y: $(a_1 + a_3 W)b_1 = a_1 b_1 + a_3 b_1 W$. Hayes (2015) named the parameter $a_3 b_1$ the *index of moderated mediation*, for this model. This index can be interpreted as follows: The indirect effect of X on Y through M is expected to differ by $a_3 b_1$ units for two cases that differ by 1 unit on the moderator (W). When the index is zero, there is no moderated mediation. Inference about this parameter provides a test of moderated mediation in this first-stage model.

Second-Stage Conditional Process Analysis. Another popular model is one which allows the relationship between M and Y to be moderated by W. Again, the direct effect can be conditional on W or not. The model for the mediator has two options: omitting or including W as a covariate:

$$\begin{aligned} M_i &= a_0 + a_1 X_i + e_{M_i} \\ M_i &= a_0 + a_1 X_i + a_2 W_i + e_{M_i} \end{aligned}$$

The effect of X on M is a_1 and is not conditional on the moderator, W. If the direct effect is moderated, the equation for the outcome is the same as that from the general model (Equation 2). If the direct effect is not moderated, the equation for the outcome is

$$Y_i = c_0' + c' X_i + (b_1 + b_3 W_i) M_i + b_2 W_i + e_{Y_i}. \tag{4}$$

In either case, the effect of M on Y is a linear function of W, $b_1 + b_3 W$. The indirect effect is the product of the effect of X on M and the effect of M on Y:

$a_1(b_1 + b_3W) = a_1b_1 + a_1b_3W$. The parameter a_1b_3 is the index of moderated mediation for this model (Hayes, 2015). The interpretation of this index is that two cases that differ by 1 unit on W are expected to differ by ab_3 on their indirect effects. Inference for this parameter provides a test of moderated mediation in the second-stage conditional process model.

Statistical Inference. Hayes (2015) showed that a null hypothesis of no moderated mediation would imply that the index of moderated mediation is zero for models where the conditional indirect effect is a linear function of the moderator. When the index of moderated mediation is zero the indirect effect is constant across the range of the moderator. Hayes (2015) recommends that inference about moderated mediation should be done using a bootstrap confidence interval, because the index of moderated mediation not always normal (Stone & Sobel, 1990). Bootstrapping is a nonparametric method which samples with replacement to generate many samples of the same size, each contributing one estimate to generate a sampling distribution (Efron & Tibshirani, 1994). Research in the area of between-subjects mediation suggests that percentile bootstrapping methods perform very well (Biesanz, Falk, & Savalei, 2010; Coutts, 2023; Hayes & Scharkow, 2013; MacKinnon, Lockwood, & Williams, 2004), and the same has been found for repeated-measures mediation (Montoya, 2022; Yzerbyt, Muller, Batailler, & Judd, 2018).

Causal Inference. While causal inference in mediation analysis has garnered much attention (Imai, Keele, & Yamamoto, 2010; Pearl, 2001; Vanderweele & Vansteelandt, 2009), the topic of causal inference for moderated mediation has been discussed much less (Qin & Wang, 2023). In the context of moderated mediation there are two potential causal effects of interest: 1) the conditional indirect effect (i.e., how much manipulation of X effects Y through M, among cases with a specific value of $W = w$), and 2) the index of moderated mediation (i.e., how much manipulation of W effects the indirect effect of X on Y through M).

In the case of the conditional indirect effect, Qin and Wang (2023) show that the typical assumptions for mediation can be applied to a subset of the population within levels of the moderator (w). Among this population, the conditional indirect effects are identified and unbiased given correct specification of the model, correct temporal ordering, and the following assumptions:

A1 No unmeasured common of X and M controlling for covariates

A2 No unmeasured common causes of X and Y controlling for covariates

A3 No unmeasured common causes of M and Y controlling for X and covariates

A4 No measured or unmeasured common causes of M and Y that are caused by X controlling for covariates

With respect to the index of moderated mediation, Loeys, Talloen, Goubrt, Moerkerke, and Vansteelandt (2016) suggest that for moderated mediation estimated with linear models which allow for W to moderate of all paths ($X - M$, $M - Y$, $X - Y$), the difference between conditional indirect effects can be more easily causally identified than indirect effects. Consider for example a moderator with two levels which is experimentally manipulated and randomly assigned. In this case, the index of moderated mediation is an unbiased estimate of the causal effect of that manipulation on the indirect effect, regardless of whether there is or is not confounding of the X, M, and Y relationships. So if the index of moderated mediation is non-zero, it is impossible for all the indirect effects to be zero. This suggests that by identification of the index of moderated mediation, a claim of mediation can be supported without evidence for A1 - A4. However, for any paths which are not moderated by a randomly assigned W, we must assume that path is not confounded. For example, in a first-stage conditional process analysis, there is no moderation of the $M - Y$ path, so for the index of moderated mediation to be identified this path must not be confounded (A3, i.e., b_1 is an unbiased estimate of the causal effect of M on Y) . In a case where the moderator is not manipulated, Loeys et al. (2016) outline the following set of assumptions to identify the index of moderated mediation:

A5 All confounders have an additive effect on M and Y

A6 No confounders interact with any other predictors of the model

A7 The linear association between predictors and confounders is the same for all levels of the moderator

Causal inference for moderated mediation is still a relatively new area; however, sets of assumptions have been described which allow for identification of important causal estimands in moderated mediation (Loeys et al., 2016; Qin & Wang, 2023). It is important for researchers to understand that if these assumptions are not met, the estimates from the statistical models do not reflect the causal effects of interest. In the case when assumptions may not be met, sensitivity analyses can be used to evaluate the impact of assumption violations on estimates from a specific dataset (Qin & Wang, 2023).

Mediation and Moderation in Two-Instance Repeated-Measures Designs

A two-instance repeated-measures design can manifest in many different ways. Participants can respond under two different conditions with order randomized, called AB/BA cross-over designs. For example, touring both of two computer science classrooms (one stereotypical and one neurtral), with the order of the classroom randomly assigned (Cheryan, Meltzoff, & Kim, 2011). Another example of a two-instance repeated-measures design is a pre-post design, where participants are measured before and after some kind of treatment or intervention. When describing two-instance repeated-measures designs the word “instance" is used, rather than condition, to capture both types of designs.

Mediation. For mediation analysis with a single mediator in a two-instance repeated-measures design, each individual (i) has four measurements: two measurements of the outcome (Y_{1i} and Y_{2i}) and two measurements of the mediator (M_{1i} and M_{2i}). N denotes the number of individuals in the study, so that $i = 1, \ldots, N$.

The estimate of interest in this analysis is the indirect effect of instance on the outcome through the mediator. It is calculated as the product of the effect of instance on the mediator and the effect of the mediator on the outcome controlling for instance. The first of these can be defined using an intercept-only model for each of the mediators.

$$M_{1i} = a_1 + e_{M_{1i}} \tag{5}$$

$$M_{2i} = a_2 + e_{M_{2i}} \tag{6}$$

Each mediator has its own mean a_1 and a_2 respectively for Instance 1 and Instance 2. The errors are assumed to be normally distributed with mean of zero and variance covariance matrix Σ_M. Taking the difference between equations for each instance results in a single equation which includes the parameter of interest: a, the effect of instance on the mediator.

$$M_{2i} - M_{1i} = (a_2 - a_1) + (e_{M_{2i}} - e_{M_{1i}}) = a + e_{M_i} \tag{7}$$

To quantify the indirect effect of instance on the outcome requires a parameter for the effect of the mediator on the outcome. Judd et al. (2001) use the following models where the outcome in Instance 1 is a linear function of the mediator in Instance 1. Similarly, the outcome in Instance 2 is a linear function of the mediator in Instance 2.

$$Y_{1i} = c_1'^{*} + b_{11} M_{1i} + e_{Y_{1i}} \tag{8}$$

$$Y_{2i} = c_2'^{*} + b_{21} M_{2i} + e_{Y_{2i}} \tag{9}$$

Similar to the previous equations, $c_1'^{*}$ and $c_2'^{*}$ are intercepts where the subscript denotes instance. The slopes b_{11} and b_{21} indicate the relationship between the mediator and the outcome in their respective instances denoted by the first subscript. The errors $e_{Y_{1i}}$ and $e_{Y_{2i}}$ are assumed to be normally distributed with means of zero and variance covariance matrix Σ_Y.

In these models there are two parameters that estimate the relationship between the mediator and the outcome variable, b_{11} and b_{21}. Judd et al. (2001) suggest using the average of these two parameters to represent the relationship between the mediator and the outcome. To estimate the average of the coefficients, the difference between Equation 8 and

Equation 9 is taken first, then the difference and average of the mediators are substituted into the model replacing the individual mediators[1]. Finally, the average of the mediators is grand mean centered to improve interpretability of the intercept (for a more in-depth explanation see Judd et al. (2001) and Montoya and Hayes (2017)):

$$\begin{aligned} Y_{2i} - Y_{1i} &= c_2'^* - c_1'^* + b_{21}M_{2i} - b_{11}M_{1i} + e_{Y_{2i}} - e_{Y_{1i}} \\ &= c' + (b_{21} - b_{11})\overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}} + \frac{b_{21} + b_{11}}{2}(M_{2i} - M_{1i}) + \\ &\quad (b_{21} - b_{11})(\frac{M_{2i} + M_{1i}}{2} - \overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}}) + e_{Y_{2i}} - e_{Y_{1i}} \end{aligned} \tag{10}$$

Substituting notation to simplify,

$$\begin{aligned} c' = c_2'^* - c_1'^* + d\overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}} \\ b = \frac{b_{21} + b_{11}}{2} \\ d = b_{21} - b_{11} \\ e_{Y_i} = e_{Y_{2i}} - e_{Y_{1i}} \end{aligned}$$

leads to a simplified equation,

$$Y_{2i} - Y_{1i} = c' + b(M_{2i} - M_{1i}) + d(\frac{M_{2i} + M_{1i}}{2} - \overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}}) + e_{Y_i} \tag{11}$$

The difference between the outcomes is modeled as a linear function of the difference between the mediators and the mean-centered average of the mediators. The weights of these predictors align with the parameters of primary interest. Specifically, $b = \frac{b_{21}+b_{11}}{2}$ is the effect of the mediator on the outcome averaged over the two instances. The coefficient $d = (b_{21} - b_{11})$ is the difference in the relationship between the mediator and the outcome across instances. The intercept of this equation c' is the difference in the outcomes for individuals with no difference in the mediators ($M_{2i} - M_{1i} = 0$) and for whom the average of the mediators is at the sample average ($(M_{2i} + M_{1i})/2 = \overline{\frac{M_{2\cdot}+M_{1\cdot}}{2}}$). The coefficient c' is the direct effect evaluated at mean of the average of the mediators.

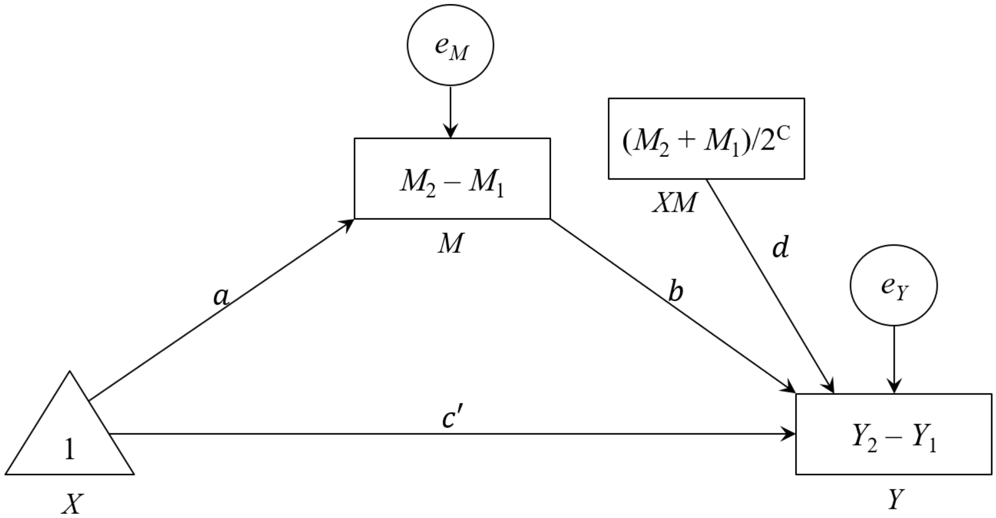

Figure 1. Path Diagram for Mediation Model for Two-Instance Repeated-Measures Design. ^{C} indicates a variable is grand-mean centered. Text inside boxes indicates statistical variables in analysis. Text below boxes indicates conceptual role of each variable.

All of these equations can be represented in a path diagram (Figure 1), with intercepts represented in the paths from the triangle to the outcome variable in the model. The indirect effect is the product of a and b and c' is the direct effect. Equations 7 and 10 can be estimated separately using ordinary least squares (OLS) regression analysis or simultaneously using a structural equation model (SEM).

There are two important assumptions being made as part of Equations 5, 6, 8 and 9. The first is there are no autoregressive effects: levels of a variable (i.e., Y_1, M_1) do not predict other levels of the same variable (i.e., Y_2 and M_2, respectively). It is important to note that this does not imply that these measures are independent, as the models do allow for covariance of the residuals. A second assumption in Equations 8 and 9 is that there are

[1] This approach is commonly used to compare regression coefficients and is described more in-depth in Darlington and Hayes (2017, pg. 231)

no crosslag relationships. This means the mediator from one instance, does not predict the outcome in the other instance above and beyond the mediator from the same instance. For example, in a model of the outcome from Instance 1 (Y_1), the mediator from Instance 2 (M_2) does not predict this outcome above and beyond the mediator from Instance 1 (M_1). The reverse is also the case: M_1 does not predict Y_2 above and beyond M_2.

Statistical Inference. Inference for the indirect effect can be conducted in a variety of ways. Judd et al. (2001) recommended a piecemeal test, such that evidence of mediation would be established if multiple coefficients are significantly different from zero, but this method does not perform well (Montoya, 2022; Yzerbyt et al., 2018). Instead, percentile bootstrap confidence intervals, the joint significance test, and Monte Carlo confidence intervals all perform well for inference on the indirect effect in repeated-measures designs (Montoya, 2022; Yzerbyt et al., 2018). MEMORE, a free and easy-to-use macro for SPSS and SAS, can conduct inference for indirect effects in two-instance repeated-measures designs using either bootstrapping or the Monte Carlo confidence interval (Montoya & Hayes, 2017, downloadable from akmontoya.com).

Causal Inference. To discuss causal inference for indirect effects in repeated-measures designs, it is important to differentiate between pre-post designs and AB/BA crossover designs. For both designs, the assumption that there are no carry-over effects is required to identify the indirect effect (Josephy, Vansteelandt, Vanderhasselt, & Loeys, 2015). The case of pre-post designs is much more similar to a cross-sectional design, where the design of the study does not limit confounding of any paths in the mediation. This means that the primary assumptions for mediation (A1 - A4) as described in between-subject designs are required for these designs as well.

However, similar to the case of a between-subject experiment, AB/BA designs do not require the assumption that there are no omitted confounders for $X - M$ (A1) and $X - Y$ (A2) (Josephy et al., 2015). A3 and A4 are still needed, except that Josephy et al. (2015) showed that if unmeasured upper-level confounders of the $M - Y$ relationship have an additive effect and there is no unmeasured heterogeneity in $X - M$, $M - Y$, and $X - Y$ effects, then upper-level confounders can be ignored for these designs. For a comparison of between-subject and repeated-measures designs on causal assumptions for indirect effects, see (Montoya, 2022).

Moderation. For two-instance repeated-measures designs, Judd et al. (1996) and Judd et al. (2001) described methods for estimating and testing if the effect of instance on some outcome is moderated by (or conditional on) a person-level covariate. This procedure begins with a model of the outcome variable Y predicted by W, the between-participant moderator, where the regression weights for this model are allowed to vary by instance (denoted in the first subscript). There are two models, one for each outcome variable:

$$Y_{1i} = b_{10} + b_{11}W_i + e_{1i}$$
$$Y_{2i} = b_{20} + b_{21}W_i + e_{2i}$$

Each outcome Y_{1i} and Y_{2i} is a linear function of W_i. The terms e_{1i} and e_{2i} are the error terms and are assumed to be normally distributed with means of zero and covariance matrix Σ_e. The intercept and slope are allowed to differ by instance. The coefficients b_{11} and b_{21} represent the relationship between W and Y in the first instance and second instance, respectively. The difference between these two regression equations is taken so that the outcome $Y_{2i} - Y_{1i}$ is the effect of instance on the outcome for each individual, and W is used as a predictor of this difference. Conceptually if W predicts the effect of instance on the outcome (i.e., $Y_{2i} - Y_{1i}$) then the effect of instance on the outcome varies across levels of W.

$$Y_{2i} - Y_{1i} = b_{20} - b_{10} + (b_{21} - b_{11})W_i + (e_{2i} - e_{1i})$$
$$Y_{2i} - Y_{1i} = b_0 + b_1W_i + e_i \quad (12)$$

This means that if the b_1 coefficient is non-zero in the population, W moderates the effect of instance on the outcome. Notice that $b_1 = b_{11} - b_{21}$, so when $b_1 = 0$ then $b_{11} = b_{21}$, meaning the relationship between W and Y is constant across instances. This matches the symmetry property observed in between-subject moderation as well: When W moderates the effect of X on Y, then X moderates the effect of W on Y.

Conducting a hypothesis test on $\hat{b}_1$, is a test of interaction between W and instance (The “hat" notation is used to represent a sample estimate of a population parameter). If $\hat{b}_1$ is significantly different from zero, the null hypothesis that b_1 is zero is rejected, suggesting that the relationship between instance and Y depends on W.

Given evidence of an interaction, additional questions arise: What is the effect of instance for those high on W, low on W, average on W? Are these effects statistically distinguishable from zero? These questions can be answered by probing an interaction using the simple-slopes procedure as described in Montoya (2019). The ratio of the estimate of the conditional effect of X on Y to its standard error is t-distributed with $N - q - 1$ degrees of freedom, where N is the number of observations and q is the number of predictors in the regression model. In the case of Equation 12, $q = 1$. Specific values of W can be plugged into the equations for the conditional effect and its standard error. The ratio of these two values gives a t-statistic, and from that a p-value can be calculated. Similarly, the conditional effect of W on Y in each instance can be calculated and tested. MEMORE conducts these calculations automatically when fitting any moderation model (Montoya, 2019).

It is often not enough to know whether or not there is an interaction, but also researchers need to know what the effect of instance is at different levels of the moderator in order to interpret the interaction. This continues to be the case in conditional process analysis, where probing will allow estimation of the indirect effect at different values of the moderator, not just testing whether or not there is moderated mediation.

Conditional Process Analysis for Two-Instance Repeated-Measures Designs

In this section, I provide a general model for estimating conditional process models in two-instance repeated-measures designs with one moderator and one mediator. Then, I will describe how simplifications of this general model correspond to more commonly used conditional process models, such as first-stage and second-stage conditional process models. For these models, I will derive the conditional indirect effect and index of moderated mediation and describe how to conduct inference useful for assessing hypotheses about moderated mediation.

A General Conditional Process Model

In the most general model with a single mediator M and a single moderator W, the moderator would be allowed to moderate four paths: the effect of instance on the mediator, the direct effect of instance on the outcome, the effect of the mediator on the outcome, and the degree to which instance moderates the relationship between M and Y. First I describe how to allow the effect of instance on the mediator, M, to be conditional on W. This is identical to the case of moderation in a two-instance repeated-measures design, where the outcome is the mediator.

Writing out the model for each mediator in each instance, the moderator W is a predictor of the mediator in each instance. When the difference between these two models is taken, the coefficient for W reflects how much the effect of instance depends on W.

$$M_{1i} = a_{10} + a_{11}W_i + e_{M_{1i}} \tag{13}$$

$$M_{2i} = a_{20} + a_{21}W_i + e_{M_{2i}} \tag{14}$$

$$M_{2i} - M_{1i} = a_{20} - a_{10} + (a_{21} - a_{11})W_i + e_{M_{2i}} - e_{M_{1i}}$$

$$M_{2i} - M_{1i} = a_0 + a_1 W_i + e_{M_i} \tag{15}$$

The effect of instance, as represented by the difference score on the mediators, is a function of the moderator W. When a_1 is different from zero, W predicts the effect of instance on the mediator, meaning this effect is conditional on W. When a_1 is zero, the effect of instance on the mediator is constant across the range of the moderator (i.e., not moderated).

Next, the model of the outcome can allow the direct effect of instance to be conditional on W. Consider the original equations from the mediation models (Equations 8 and 9) but adding W as a predictor with differing slopes by instance. Then the difference is taken and the same substitution of the mediators is applied.

$$\begin{aligned}
Y_{1i} &= c_1'^* + b_{11}M_{1i} + b_{12}W_i + e_{Y_{1i}} \\
Y_{2i} &= c_2'^* + b_{21}M_{2i} + b_{22}W_i + e_{Y_{2i}} \\
Y_{2i} - Y_{1i} &= (c_2'^* - c_1'^*) + b_{21}M_{2i} - b_{11}M_{1i} + (b_{22} - b_{12})W_i + (e_{Y_{2i}} - e_{Y_{1i}}) \\
Y_{2i} - Y_{1i} &= (c_2'^* - c_1'^*) + \frac{b_{21} + b_{11}}{2}(M_{2i} - M_{1i}) + (b_{21} - b_{11})\frac{M_{2i} + M_{1i}}{2} \\
&\quad + (b_{22} - b_{12})W_i + (e_{Y_{2i}} - e_{Y_{1i}}) \qquad (16)
\end{aligned}$$

If the relationship between W and Y is the same across instances (i.e., $b_{12} = b_{22}$) then the coefficient for W in Equation 16 will be zero. This would mean that W does not predict the effect of instance on Y after controlling for the mediator. By allowing W to be a predictor in the equation for each outcome with different weights, the direct effect of instance on Y is conditional on W. Thus estimating Equation 16, the direct effect is allowed to be moderated by W.

The next step allows the relationship between M and Y to depend on W. Building on the previous model, consider a case where the degree to which the relationship between M and Y depends on W is fixed to be constant across instance.

$$Y_{1i} = c_1'^* + (b_{11} + b_{.3}W_i)M_{1i} + b_{12}W_i + e_{Y_{1i}}$$

$$Y_{2i} = c_2'^* + (b_{21} + b_{.3}W_i)M_{2i} + b_{22}W_i + e_{Y_{2i}}$$

The relationship between M and Y when W is zero is allowed to differ by instances (i.e., b_{11} is not necessarily equal to b_{21}). However, the degree to which the relationship between M and Y varies as W varies, $b_{.3}$, is fixed across instances. This means that the degree to which W moderates the relationship between M and Y is not itself moderated by instance. The between-subject analog would be including three two-way interactions (XW, XM, and MW) in the model for Y, but not a three-way interaction XMW.

By taking the difference between the two equations and grouping like terms, the result is

$$\begin{aligned} Y_{2i} - Y_{1i} &= (c_2'^* - c_1'^*) + (b_{21} + b_{\cdot 3}W_i)M_{2i} - (b_{11} + b_{\cdot 3}W_i)M_{1i} + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}} \\ &= (c_2'^* - c_1'^*) + b_{21}M_{2i} - b_{11}M_{1i} + b_{\cdot 3}W_i(M_{2i} - M_{1i}) + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}} \end{aligned}$$

Using the same substitution applied in the mediation model previously:

$$\begin{aligned} Y_{2i} - Y_{1i} = (c_2'^* - c_1'^*) + \frac{b_{21} + b_{11}}{2}(M_{2i} - M_{1i}) + (b_{21} - b_{11})\frac{M_{2i} + M_{1i}}{2} \\ + b_{\cdot 3}W_i(M_{2i} - M_{1i}) + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}} \end{aligned} \quad (17)$$

Including the product of the moderator and the difference between the mediators $W_i(M_{2i} - M_{1i})$ allows the M to Y relationship to be moderated by W. The coefficient for the difference between the mediators $\frac{b_{21}+b_{11}}{2}$ is the effect of M on Y when W is zero, averaged across instance. The coefficient for the average of the mediators $b_{21} - b_{11}$ is the difference between the effect of M on Y controlling for W. The coefficient for the product of the moderator and the difference in the mediators $b_{\cdot 3}$ is the difference in the relationship between M and Y when W differs by one unit. This coefficient, when it is different from zero reflects moderation of the effect of M on Y by W. As before, the coefficient for W, $b_{22} - b_{21}$, reflects the degree to which the direct effect is moderated by W.

Next, consider a model where the moderation of the effect of M and Y is allowed to differ by instance. This is equivalent to allowing a three-way interaction between M, W, and instance. This is the most general model discussed in this manuscript. Beginning with the same model as above, but allowing the coefficient for the product between M and W to differ by instance:

$$Y_{1i} = c_1'^* + (b_{11} + b_{13}W_i)M_{1i} + b_{12}W_i + e_{Y_{1i}} \quad (18)$$

$$Y_{2i} = c_2'^* + (b_{21} + b_{23}W_i)M_{2i} + b_{22}W_i + e_{Y_{2i}} \quad (19)$$

The relationship between M and Y when W is zero is allowed to differ by instance, and the degree to which the relationship between M and Y varies as W varies is allowed to differ by instance. Described another way, the degree to which W moderates the relationship between M and Y is moderated by instance, or the degree to which instance moderates the relationship between M and Y is moderated by W (a three-way interaction).

Taking the difference between the two equations, grouping like terms, and applying the substitution results in:

$$\begin{aligned}
Y_{2i} - Y_{1i} &= (c_2'^{*} - c_1'^{*}) + (b_{21} + b_{23}W_i)M_{2i} - (b_{11} + b_{13}W_i)M_{1i} + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}} \\
&= (c_2'^{*} - c_1'^{*}) + \frac{b_{21} + b_{23}W_i + b_{11} + b_{13}W_i}{2}(M_{2i} - M_{1i}) \\
&\quad + (b_{21} + b_{23}W_i - b_{11} - b_{13}W_i)\frac{M_{2i} + M_{1i}}{2} + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}} \\
&= (c_2'^{*} - c_1'^{*}) + \frac{b_{21} + b_{11}}{2}(M_{2i} - M_{1i}) + \frac{b_{23} + b_{13}}{2}W_i(M_{2i} - M_{1i}) + (b_{21} - b_{11})\frac{M_{2i} + M_{1i}}{2} \\
&\quad + (b_{23} - b_{13})W_i\frac{M_{2i} + M_{1i}}{2} + (b_{22} - b_{12})W_i + e_{Y_{2i}} - e_{Y_{1i}}. \qquad (20)
\end{aligned}$$

This equation differs from Equation 17 by including an additional term which is the product of the moderator and the average of the mediators. The coefficient for this term $(b_{23} - b_{13})$ is the difference in the degree to which W moderates the relationship between M and Y in each instance. This tells us if the moderation of the M–Y relationship by W is moderated by instance. The interpretation of all other parameters stays the same with one exception: Previously, the coefficient for the product of the moderator and the difference of the mediators estimated the degree to which the relationship between M and Y varies as W varies (in both instances). Now the coefficient is an estimate of the degree to which the relationship between M and Y varies as W varies, averaged over instance.

Just as before, the average of the mediators is grand mean-centered, and Equation 20 can be rewritten as

$$\begin{aligned}
Y_{2i} - Y_{1i} = c_0' + c_1'W_i + b_1(M_{2i} - M_{1i}) + d_1\left(\frac{M_{2i} + M_{1i}}{2} - \overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}}\right) \\
+ b_2W_i(M_{2i} - M_{1i}) + d_2W_i\left(\frac{M_{2i} + M_{1i}}{2} - \overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}}\right) + e_{Y_i} \qquad (21)
\end{aligned}$$

where

$$c_0' = c_2'^{*} - c_1'^{*} + b_2\overline{\frac{M_{2\cdot} + M_{1\cdot}}{2}}$$

$$\begin{aligned}
c_1' &= b_{22} - b_{12} + (b_{23} - b_{13})\frac{\overline{M_{2.} + M_{1.}}}{2} \\
b_1 &= \frac{b_{21} + b_{11}}{2} \\
d_1 &= b_{21} - b_{11} \\
b_2 &= \frac{b_{23} + b_{13}}{2} \\
d_2 &= b_{23} - b_{13} \\
e_{Y_i} &= e_{Y_{2i}} - e_{Y_{1i}}
\end{aligned}$$

In combination, Equation 15 and 21 represent the most general model discussed in this paper, which allows W to moderate four effects: the effect of instance on the mediator, the direct effect of instance on the outcome, the effect of the mediator on the outcome, and the degree to which instance moderates the relationship between M and Y. This set of equations is represented both statistically, as a path diagram, in Figure 2 (top) and conceptually, representing the conditional effects in Figure 2 (bottom). Table 1 provides an interpretation for each coefficient and conditional effect in this general model. Note that these interpretations are sensitive to the scaling and coding of the variables in the model. For example, many coefficients are conditional on variables being at a value of zero or a one-unit difference on a variable. Researchers should be mindful of this issue and avoid interpreting coefficients which may result in extrapolation beyond the range of the observed data. Standardizing and mean centering mediators and outcomes is not recommended as their mean structure must be maintained for accurate estimation of the model.

Table 1

Coefficient Interpretations for General Moderated Mediation Model

Coefficient	Interpretation
$\hat{a}_0$	For cases with a value of zero on the moderator, the mediator is expected to be $\hat{a}_0$ units higher in Instance 2 compared to Instance 1.
$\hat{a}_1$	For two cases that differ by 1 unit on the moderator, the effect of instance on the mediator is expected to be $\hat{a}_1$ units higher for the case with the higher value on the moderator.
$\hat{c'}_0$	For cases that show no difference on the mediator measures, are average on the mediators, and score zero on the moderator, the outcome variable is expected to be $\hat{c'}_0$ units higher in Instance 2 compared to Instance 1.
$\hat{c'}_1$	For two cases that show no difference on the mediator measures and are average on the mediators, but differ by one unit on the moderator, the effect of instance on the outcome is expected to be $\hat{c'}_1$ units higher for the case with the higher value on the moderator.

Coefficient Interpretations for General Moderated Mediation Model (continued)

Coefficient	Interpretation
$\hat{b}_1$	Option 1: For two cases with a value of zero on the moderator, the same value on the average of the mediators, and differ by 1 unit on the difference between the mediators (i.e., the effect of instance on the mediators), the difference between the outcomes is expected to be $\hat{b}_1$ units higher for the case with the higher difference between the mediators. Option 2: Averaging over Instance 1 and Instance 2, for two cases that score zero on the moderator and differ by one unit on the mediator, they are expected to differ by $\hat{b}_1$ units on the outcome.
$\hat{d}_1$	Option 1: For two cases with a balue of zero on the moderator, the same value on the difference between the mediators, and differ by 1 unit on the average between the mediators, the difference between the outcomes is expected to be $\hat{d}_1$ units higher for the case with the higher average between the mediators. Option 2: For cases that score zero on the moderator, how much a one unit difference on the mediator results in expected differences on the outcome, is $\hat{d}_1$ greater in Instance 1 compared to Instance 2.

Coefficient Interpretations for General Moderated Mediation Model (continued)

Coefficient	Interpretation
$\hat{b}_2$	Option 1: For two cases that differ by one unit on the moderator, a 1 unit difference on the difference between the mediators (i.e., the effect of instance on the mediators) is expected to result in a $\hat{b}_2$ larger difference between the outcomes among the case with the higher moderator. Option 2: Averaging over Instance 1 and Instance 2, for two cases that differ by one unit on the moderator, a 1 unit difference on the mediator is expected to result in a $\hat{b}_2$ larger difference on the outcome among the case with the higher value on the moderator.
$\hat{d}_2$	Option 1: For two cases that differ by one unit on the moderator, a 1 unit difference on the average of the mediators is expected to result in a $\hat{d}_2$ larger difference between the outcomes, among the case with the higher moderator. Option 2: For two cases that differ by one unit on the moderator, the difference between how much a one unit difference on the mediator results in expected differences on the outcome in Instance 1 compared to Instance 2, is expected to be $\hat{d}_1$ greater in the case with the higher value on the moderator.
$\hat{\theta}_{X \to M}(W = w) = \hat{a}_0 + \hat{a}_1 w$	For cases with a value of w on the moderator, the mediator is expected to be $\hat{\theta}_{X \to M}(W = w)$ units higher in Instance 2 compared to Instance 1.

Coefficient Interpretations for General Moderated Mediation Model (continued)

Coefficient	Interpretation
$\hat{\theta}_{M\to Y}(W = w) = \hat{b}_1 + \hat{b}_2 w$	For two cases with a value of w on the moderator and differ by 1 unit on the difference between the mediators (i.e., the effect of instance on the mediators), the difference between the outcomes is expected to be $\hat{\theta}_{M\to Y}(W = w)$ units higher for the case with the higher difference between the mediators. Option 2: Averaging over Instance 1 and Instance 2, for two cases with a value of w on the moderator and differ by one unit on the mediator, they are expected to differ by $\hat{\theta}_{M\to Y}(W = w)$ units on the outcome.
$\hat{\theta}_{X\to Y}(M_A = \overline{M_A}, W = w) = \hat{c'}_0 + \hat{c'}_1 w$	For cases with no difference on the mediators, are average on the mediators, and score w on the moderator, the outcome variable is expected to be $\hat{\theta}_{X\to Y}(M_A = \overline{M_A}, W = w)$ units higher in Instance 2 compared to Instance 1.
$\hat{\theta}_{X\to M\to Y}(W = w) = (\hat{a}_0 + \hat{a}_1 w)(\hat{b}_1 + \hat{b}_2 w)$	For cases with a value of w on the moderator, the outcome is expected to be $\hat{\theta}_{X\to M\to Y}(W = w)$ units higher in Instance 1 compared to Instance 2 indirectly through the effect of instance on the mediator and the mediator's subsequent effect on the outcome.

Note: For coefficients with multiple possible interpretations, Options 1 and 2 are presented. Both interpretations are accurate and correct, though one may be more understandable in a given context.

Defining Conditional Effects. Equations 15 and 21 can be used to define the important conditional effects for assessing moderation of mediation. These conditional effects will be used as the components that make up the conditional indirect effect, conditional direct effect, and the index of moderated mediation.

Equation 15 gives the estimate of the conditional effect of instance on the mediator:

$$\theta_{X\to M}(W) = a_0 + a_1 W \tag{22}$$

This is how much the mediators in each instance are expected to differ given a specific value of the moderator, W. Next, by grouping terms in Equation 21 based on the difference and mean-centered average of the mediators the result is

$$Y_{2i} - Y_{1i} = c_0' + c_1' W_i + (b_1 + b_2 W_i)(M_{2i} - M_{1i}) + (d_1 + d_2 W_i)\left(\frac{M_{2i} + M_{1i}}{2} - \frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}\right) + e_{Y_i} \tag{23}$$

The coefficient for the difference between the mediators is $\theta_{M\to Y}(W) = b_1 + b_2 W$ which estimates the conditional effect of M on Y averaged over instance. The conditional indirect effect is of primary interest when considering hypotheses of moderated mediation. The conditional indirect effect is the product of the effect of instance on the mediator and the effect of the mediator on the outcome averaged over instance. In the most general model, each of these effects are linear functions of W. As such the indirect effect will be conditional on W.

$$\begin{aligned}\theta_{X\to M\to Y}(W) &= \theta_{X\to M}(W)\theta_{M\to Y}(W) \\ &= (a_0 + a_1 W)(b_1 + b_2 W) \\ &= a_0 b_1 + (a_0 b_2 + a_1 b_1)W + a_1 b_2 W^2\end{aligned}$$

Unlike the other conditional effects in this model, the conditional indirect is not a linear function of W. Rather, it is a quadratic function of W. When estimating this model, specific values of W may be selected and the conditional indirect effect can be estimated. An index of moderated mediation cannot be calculated in this general model, because there is no single coefficient that represents the *linear* relationship between the moderator and the conditional indirect effect. However, for a dichotomous moderator the two conditional indirect effects can be estimated and their difference can be used to test moderated mediation (following the approach described by Fairchild and MacKinnon (2009)).

The conditional direct effect is the expected difference in outcomes for an individual who shows no difference on the mediators. By taking Equation 23 and setting $M_{2i} - M_{1i} = 0$ the following equation results

$$E(Y_{2i}-Y_{1i}|M_{2i}-M_{1i}=0) = c'_0+c'_1W_i+d_1(\frac{M_{2i}+M_{1i}}{2}-\frac{\overline{M_{2\cdot}+M_{1\cdot}}}{2})+d_2W_i(\frac{M_{2i}+M_{1i}}{2}-\frac{\overline{M_{2\cdot}+M_{1\cdot}}}{2}) \tag{24}$$

The direct effect depends both on the average of the mediators (as it did in the simple mediation case) and moderator W. In line with the simple mediation case, conditioning on the mean of the average of the mediators gives:

$$\theta_{X\rightarrow Y}(M_A = \overline{M_A}, W) = c'_0 + c'_1W \tag{25}$$

Here $M_{Ai} = \frac{M_{2i}+M_{1i}}{2}$ and $\overline{M_A} = \frac{\overline{M_{2\cdot}+M_{1\cdot}}}{2}$. For two cases who are average on the mediators, but differ by one unit on W the direct effect is expected to be c'_1 units higher for the case with the higher value of W.

Allowing each path in the mediation pathway to be moderated is very general, but it lacks parsimony. If only certain paths are hypothesized to be moderated, then specifying a more parsimonious model would be useful. Using the general model proposed above, I now show how to constrain the model to give a variety of more popular conditional process models.

First-Stage Conditional Process Models

In a first-stage conditional process model, W is allowed to moderate the relationship between instance and the mediator. It is not allowed to moderate the relationship between the mediator and the outcome. I will discuss cases when the direct effect is and is not moderated. Because the relationship between instance and the mediator is moderated, the equation for the difference in mediators is the same as Equation 15 and the conditional effect of instance on the mediators is still represented by Equation 22.

The relationship between M and Y is not allowed to vary across the moderator. When the direct effect is allowed to be moderated by W, applying the substitution to

Equation 16 and centering the average of the mediators results in

$$\begin{aligned} Y_{2i} - Y_{1i} &= (c_2'^{*} - c_1'^{*}) + (b_{21} - b_{11})(\frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}) + \frac{b_{21} + b_{11}}{2}(M_{2i} - M_{1i}) \\ &+ (b_{21} - b_{11})(\frac{M_{2i} + M_{1i}}{2} - \frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}) + (b_{22} - b_{12})W_i \\ &+ (e_{Y_{2i}} - e_{Y_{1i}}) \end{aligned} \quad (26)$$

with simplified notation

$$Y_{2i} - Y_{1i} = c_0' + c_1'W_i + b_1(M_{2i} - M_{1i}) + d_1(\frac{M_{2i} + M_{1i}}{2} - \frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}) + e_{Y_i} \quad (27)$$

The effect of the mediator on the outcome, averaging over instance is:

$$\theta_{M \to Y} = (b_{21} + b_{11})/2 = b_1 \quad (28)$$

By constraining b_2 and d_2 to be zero in the general model (Equation 21), the result is a first-stage conditional process model, which allows the direct effect to be moderated.

To also remove moderation of the direct effect, c_1' can be set to 0 which would result Equation 10, which was used for estimating the mediation model. The effect of instance on the mediators is allowed to be moderated by W, and so the product of this conditional effect of instance on the mediator and the unconditional effect of the mediator on the outcome will give the following conditional indirect effect:

$$\theta_{X \to M \to Y}(W) = \theta_{X \to M}(W)\theta_{M \to Y} = (a_0 + a_1W)b_1 = a_0b_1 + a_1b_1W \quad (29)$$

The conditional indirect effect is a linear function of W and a test on a_1b_1 will provide a test of whether the indirect effect is moderated by W. Here a_1b_1 is the index of moderated mediation (Hayes, 2015), which can be interpreted as the expected difference in the indirect effect for two cases that differ by one unit on W. Note that the coefficient used to estimate the effect of the mediators on the outcome does not depend on whether or not the direct effect is allowed to be moderated, it is b_1. But, whether or not the direct effect is allowed to be moderated will impact the *estimate* of b_1 ($\hat{b}_1$).

Second-Stage Conditional Process Model

In a second-stage conditional process model, the moderator does not moderate the effect of instance on the mediator, but it does moderate the effect of the mediator on the outcome. Like a first-stage conditional process model, the direct effect can be moderated or not, and there can be moderation of the d paths or not. The relationship between W and the mediator is fixed to be the same across instance, thus fixing the effect of instance on the mediator across levels of the moderator.

$$M_{1i} = a_{10} + a_{.1}W_i + e_{M_{1i}}$$
$$M_{2i} = a_{20} + a_{.1}W_i + e_{M_{2i}}$$

When the difference between these two equations is taken, the W terms cancel out, the same model which was used for the difference in the mediators when doing simple mediation analysis (Equation 7).

$$\theta_{X\to M} = a_0$$

For the model of the outcome, consider the case where both the relationship between the mediator and the outcome and the direct effect are moderated. This is represented by Equation 21. Thus the conditional effect of M on Y is $\theta_{M\to Y}(W) = b_1 + b_2W$. If the direct effect is not hypothesized to be moderated, c'_1 can be constrained to 0, meaning the W would not be included in the model on it's own, but only in the products with the difference in and average of the mediators.

$$\begin{aligned} Y_{2i} - Y_{1i} &= c'_0 + b_1(M_{2i} - M_{1i}) + d_1(\frac{M_{2i} + M_{1i}}{2} - \frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}) \\ &\quad + b_2W_i(M_{2i} - M_{1i}) + d_2W_i(\frac{M_{2i} + M_{1i}}{2} - \frac{\overline{M_{2\cdot} + M_{1\cdot}}}{2}) + e_{Y_i} \end{aligned} \tag{30}$$

Estimating the conditional indirect effect will proceed similarly regardless of whether the direct effect is moderated. However, including W in the model will impact the estimates of b_1 and b_2. Care should be taken to consider whether theory predicts a moderated direct effect or not, and to include W in the model in the appropriate places (Fossum, 2023). The product of the effect of instance on the mediator and the conditional effect of the mediator on the outcome is the conditional indirect effect.

$$\theta_{X \to M \to Y}(W) = \theta_{X \to M} \theta_{M \to Y}(W) = a_0(b_1 + b_2 W) = a_0 b_1 + a_0 b_2 W \tag{31}$$

Again the conditional indirect effect is a linear function of W and a test on the index of moderated mediation, $a_0 b_2$, will provide a test of whether the indirect effect is moderated by W. This index can be interpreted as the expected difference in conditional indirect effects for two cases that differ by one unit on W.

Statistical Inference for Moderated Mediation

There are two primary parameters relevant to hypotheses about moderated mediation: the index of moderated mediation and conditional indirect effects. Bootstrapping methods can be used to generate confidence intervals for each coefficient and combinations of coefficients. Table 2 defines the conditional indirect effect for first-stage, second-stage, and the general (first- and second- stage) conditional process models.

If W is dichotomous then the conditional indirect effect can be evaluated at two values of W (w_1 and w_2) and bootstrapping methods can be used to generate a confidence interval for these conditional indirect effects and their difference (Fairchild & MacKinnon, 2009). The confidence interval for the difference will provide a test of moderated mediation.

When W is continuous, the conditional indirect effect can be evaluated at any value(s) of W and bootstrapping methods can be used to generate a confidence interval for these conditional indirect effects. Table 2 provides asymptotic standard errors which could be used to generate a z-statistic, p-value, and a confidence interval; however, this method assumes that the conditional indirect effect is normally distributed. This may be true in very large samples, but not typically true in samples of the size usually collected in psychology.

The index of moderated mediation is only defined when the conditional indirect effect is a linear function of the moderator (i.e., only the effect of X on M or the effect of M on Y but not both are moderated). When the index of moderated mediation is defined, as in Table 2, a bootstrap confidence interval can be used to conduct inference on this index. If the confidence interval does not contain zero then we can be confident that the population index of moderated mediation is non-zero at a level corresponding with the confidence interval (e.g., 95%).

The next section applies the models and inferential approaches described in this section to a real data example from psychology. The computational tool MEMORE is introduced as an easy way to implement these models and provides in-depth statistical output including inference for all important estimates, probing, and plotting. Limitations to causal inference are discussed for the example as well.

Applied Examples with MEMORE

MEMORE is an easy-to-use macro for SPSS and SAS which can fit mediation (Model 1; Montoya & Hayes, 2017) and moderation models (Models 2 and 3; Montoya, 2019). Here, I introduce its use for conditional process analysis (Models 4 - 18) with a single moderator and one or more mediators in parallel. In this section, I will demonstrate the use of MEMORE with conditional process models in two-instance repeated-measures designs. This section (and a second example in Appendix D) showcase the versatility of MEMORE with respect to the different models it can fit and its available options, and the ease with which it can be used. MEMORE is documented at akmontoya.com. Figure 3 provides a subset of the available models in MEMORE, Models 4 - 18 all integrate moderated mediation into simple and parallel mediator models with one moderator. The model numbers indicate which paths in the model are allowed to be moderated, where the lower numbered models are more flexible (allow for more moderation) and the higher number more constrained (fewer paths are moderated). One applied example is presented in this section, another is provided in Appendix A. For readers interested in R implementations, Appendix B describes methods for fitting these models with the `lavaan` package.

The primary example comes from a study where all participants go through the same intervention and are measured before and after the intervention. The moderator is a continuous variable which moderates the effect of the mediator on the outcome. In Appendix A a second example includes moderation on the first stage and the direct effect. This secondary example is a pre-post design, where participants are randomized to immediate or delayed treatment. The treatment is included as a moderator of the direct effect and the path from the instance (instance) to the mediator, showing how data from a 2(within) $\times$ 2(between) can be analyzed using the conditional process models described in this paper. The approach presented in this paper treats the repeated-measures factor (time) as the focal predictor; whereas, the same design can be analyzed with the between-subject factor as the focal predictor using methods described by Valente and MacKinnon (2017). These two examples demonstrate the variety of study designs for which conditional process analysis is appropriate.

For the purposes of the examples, I use an α-level of 0.05 and 95% confidence intervals throughout. These examples were selected for pedagogical purposes and demonstrate specific patterns of significance which aid understanding of these topics. As such, they should not be used for making substantive conclusions, as they were selected to demonstrate these methods, rather than reflecting a priori hypotheses.

Applied Example: Lasselin et al., 2016

In Lasselin et al. (2016), participants experiencing chronic pain underwent a behavioral treatment to help manage pain. Participants reported their pain levels and physical quality of life prior to and after the intervention. In this example, physical quality of life is used as a potential mediator of the effect of treatment on pain levels. In Lasselin et al. (2016), baseline inflammation was identified as a moderator of the treatment effect on pain levels. This could mean that inflammation may moderate the indirect effect of treatment on pain through physical quality of life. In particular, while physical quality of life may improve through treatment, this may not translate to a decrease in pain for individuals who have high baseline inflammation. This hypothesis is reflected in Figure 4.

The proposed model can be tested using MEMORE Model 16, which allows the path from the mediator to the outcome to be moderated (i.e., the path from physical quality of life to pain levels to be moderated by baseline inflammation). Inflammation is calculated as a principle component so it has a mean close to zero and standard deviation close 1. The code to fit this model in SPSS is

```
MEMORE y = PostPain PrePain /m = PostPCS PrePCS /w = inflame /model = 16
        /seed = 1169404.
```

and the output is in the Appendix C. The results replicate the finding from Lasselin et al. (2016) that the total effect of treatment on pain is significantly moderated by baseline inflammation ($\hat{c}_1 = 0.40$, $t(76) = 3.06$, $p = .003$), such that individuals with higher levels of inflammation show smaller decreases in pain over time. The conditional effects output from MEMORE show that individuals at one standard deviation below the mean on inflammation (-1.02) show significant decreases in pain over time ($\hat{\theta}_{X\to Y}(W = -1.02) = -0.64$, $t(76) = -3.39$, $p = .001$). But for those who are average and one standard deviation above the mean on inflammation, there is not a statistically significant change in pain over time ($\hat{\theta}_{X\to Y}(W = 0) = -0.23$, $t(76) = -1.73$, $p = .09$; $\hat{\theta}_{X\to Y}(W = 1.02) = 0.179$, $t(76) = 0.95$, $p = .34$). It is important to note that effects which are not statistically significant should not be interpreted as evidence for the null hypothesis, rather an absence of evidence against the null hypothesis.

The model for the difference in the mediators (physical quality of life), shows that physical quality of life increases over time by 5.07 units on average, and this effect is statistically significant ($t(77) = 5.25$, $p < .001$). The model for the difference in the outcomes is:

$$\widehat{Y_{2i} - Y_{1i}} = -0.2429 + 0.0079(M_{2i} - M_{1i}) + 0.0334W_i(M_{2i} - M_{1i}) - 0.0211\left(\frac{M_{2i} + M_{1i}}{2}\right)^C \quad (32)$$

In the model for the difference in pain, there is not a statistically significant effect of physical quality of life when inflammation is held at zero (close to the mean; $\hat{b}_1 = .0079$, $t(74) = 0.41$, $p = .68$). But there is a statistically significant interaction between physical quality of life and inflammation such that for individuals higher in inflammation the relationship between physical quality of life and pain levels is higher than for individuals lower in inflammation ($\hat{b}_2 = .03$, $t(74) = 2.08$, $p = .04$). This is more clearly observed in the conditional effects tables in the MEMORE output: for individuals one standard deviation below the mean on inflammation, higher levels of physical quality of life correspond to lower levels of pain in this sample, but this effect is not statistically significant and cannot be generalized to the population ($\hat{\theta}_{M \to Y}(W = -1.02) = -0.03$, $t(74) = -1.48$, $p = .14$). Whereas for individuals at the mean and one standard deviation above the mean on inflammation, higher levels of physical quality of life correspond to greater levels of pain within this sample, again not statistically significant ($\hat{\theta}_{M \to Y}(W = 0) = 0.01$, $t(74) = 0.41$, $p = 68$; $\hat{\theta}_{M \to Y}(W = 1.02) = 0.04$, $t(74) = 1.35$, $p = 0.18$). A test of the coefficient for the average of physical quality of life shows no statistically significant difference in the relationship between physical quality of life and pain before and after treatment ($\hat{d}_1 = -.02$, $t(74) = -1.76$, $p = .08$). The negative coefficient means the relationship between physical quality of life and pain is more positive before compared to after treatment within this specific sample; however, this difference is not statistically significant.

Examining the conditional indirect effects, the percentile bootstrap confidence interval includes zero for individuals at the mean of inflammation ($\hat{\theta}_{X \to M \to Y}(W = 0) = 0.04$, 95% bootstrap CI = [-0.09, 0.18]). For those who are one standard deviation below the mean on inflammation there is a statistically significant negative indirect effect ($\hat{\theta}_{X \to M \to Y}(W = -1.02) = -0.13$, 95% bootstrap CI = [-0.26, -0.01]). So, for individuals who are low in inflammation at baseline, their pain is expected to decrease over treatment through the impact of treatment on physical quality of life. Specifically, treatment results in higher levels of physical quality of life, and higher levels of physical quality of life translate to lower levels of pain. Whereas, for those who are one standard deviation above the mean on inflammation, there is a statistically significant positive indirect effect ($\hat{\theta}_{X \rightarrow M \rightarrow Y}(W = 1.02) = 0.21$, 95% bootstrap CI = [0.001, 0.45]). For individuals who are high on inflammation at baseline, their pain is expected to increase over treatment through the impact of treatment on physical quality of life. Specifically, treatment results in higher levels of physical quality of life, but higher levels of physical quality of life translate to higher levels of pain.

The differing pattern of significance of the conditional indirect effects is not sufficient to support a claim that the indirect effect is moderated. For this, a test on the index of moderated mediation is needed. MEMORE provides an estimate of this index as well as a 95% bootstrap confidence interval. In this model the index of moderated mediation is $a_0 b_2$, which is estimated to be 0.17, meaning that for individuals that differ by one unit on inflammation, their indirect effects are expected to differ by 0.17 units, where those higher on inflammation are expected to have higher indirect effects. The 95% percentile bootstrap confidence interval for this coefficient does not include zero (0.06, 0.29), meaning the indirect effect is statistically significantly moderated at an α-level of 0.05. This provides clear evidence that the indirect effect of behavioral treatment on pain levels through physical quality of life is moderated by inflammation levels.

Causal Inference. While the exact mathematical identification assumptions for moderated mediation in two-instance repeated-measures designs have not been described, there are clear limitations to causal inference in this example which are worth describing. Because linear models are used to generate estimates, these models must be correctly specified for causal inference to be valid. Additionally, for all mediation analyses the correct specification of variable order is an assumption (temporal precedence). In this example, this means that inflammation should precede the change in physical quality of life, which it does because the inflammation measure is collected at baseline. Additionally, change in physical quality of life should precede change in pain. This assumption is more difficult to defend as the baseline measures of the mediator and outcome are taken during the same session, and the same is true for the measures taken after treatment. The time-scales of the questions leave room for concern as the pain measure asks participants to reflect on the “past week" specifically, whereas the physical quality of life measure asks about a "typical day". This issue of timing of measurements and how question stems can be used to aid causal inferences is returned to in the Discussion section.

Beginning with the index of moderated mediation, the example from Lasselin et al. (2016) has weak evidence for causal inference for the index of moderated mediation because the moderator is not randomly assigned. There are two parameters to consider a_0 and b_2. Starting with a_0, because there is no moderation of the X to M path, we must assume there are no confounders of this path (A1). This means that all change in physical quality of life must be assumed to be due to the treatment. There can be no natural change over time or other factors that impact physical quality of life between initiation and completion of treatment. The reasonableness of this assumption requires content expertise to evaluate. If, for example, physical quality of life has been demonstrated to be a very stable trait for individuals in the target population, change over time during the intervention could be enough to support causal claims (Shadish, Cook, & Campbell, 2002). However, if there are many demonstrated factors which affect physical quality of life that may themselves change over time, this would suggest that there is little support for a_0 as estimating a causal effect.

The b_2 parameter is a bit more difficult to assess, because inflammation is not randomly assigned. For this part of the model, we must rely on assumptions A5 – A7 as discussed by Loeys et al. (2016). Yzerbyt, Muller, and Judd (2004) note that no-confounder assumptions also apply in the context of moderation, where alternative moderators with which the candidate moderator is merely correlated (e.g., due to some common cause) must be considered. In this specific example, this would mean that perhaps there is some other variable that moderates the relationship between physical quality of life and pain levels, which is related to but does not cause and is not caused by inflammation (violating A5 and A6). Alternatively, A7 would be violated if there was some variable, for example sleep quality, that affected both physical quality of life and pain levels where inflammation moderates the effect of that confounder (sleep quality) on physical quality of life. Content expertise is required to evaluate the plausibility of these assumption. But if these assumptions were plausible, and the assumptions for a_0 were also plausible, it could be the case that the index of moderated mediation would be reasonably assumed to be causal, but only under the strict conditions described above. Future research to develop sensitivity analyses for violations of these assumptions is greatly needed.

The conditional indirect effects involve a_0, b_1, and b_2. Because a_0 and b_2 have already been discussed, the assumptions required for those effects would be required, but also for b_1 to apply, the additional no-confounder assumption is required for the relationship between physical quality of life and pain (A3 and A4). This means there can be no variables which affect both physical quality of life and pain (otherwise b_1 will be biased), but under mild assumptions of additive effects (as in Josephy et al. (2015)) only lower-level confounders (i.e., confounders that vary across time) must be considered. As in all cases, the plausibility of these assumptions must be evaluated by content experts. The role of statisticians and quantitative methodologists is to describe these assumptions in a way that content experts can reasonably evaluate in specific contexts. More work is needed to articulate the assumptions required to make causal inferences in conditional process analysis with two-instance repeated-measures designs and develop methods for sensitivity analyses.

Equivalence and Extensions to Multilevel Models

Two-instance repeated-measures designs are limited in multiple ways, and researchers may want to consider designs which are more complex. For example, collecting repeated-replicates in each instance, where instead of having one observation in each instance, there are 5 or 10. Alternatively, a pre-post design is a type of longitudinal design, but what if there are more time points? In this section, I demonstrate how the proposed model is equivalent to a 1-1-1 multilevel conditional process model model when there are no repeated-replicates, and how this model could be expanded out for repeated replicates or multiple time points.

Data from a two-instance repeated-measures design can be used to estimate conditional process analysis with a multilevel model. With one observation for each individual in each instance, some of the distinct advantages of multilevel models are not available. For example, researchers may be interested in understanding variability within an individual which is not estimable in the two-instance repeated-measures conditional process models described in this paper. However, with multiple replicates per instance, multilevel models can estimate within-individual variability. In this section, the focus is multilevel conditional process models where the mediation model is among variables measured at Level 1 (i.e., X, M, and Y are all at Level 1) and the moderator is a Level 2 variable. I show how this model is related to the models presented in the previous sections.

1-1-1 Conditional Process Models

As a slight change in notation, the instance under which the measurement was made will no longer be represented as a subscript, but rather a dichotomous predictor $X_{ik} \in 0, 1$ will represent instance. M_{ik} refers to the mediator measurement on the ith individual on the kth replicate, where $k = 1, \dots, K$ and K is the total number of measurements made on individual i rather than the number in a specific instance. The variables X_{ik} and Y_{ik} will denote the instance and outcome for the kth replicate for individual i, respectively. In a two-instance repeated-measures designs $K = 2$. The person-level mean for individual i on the mediator is $M_{i\cdot} = \frac{1}{K}\sum_{k=1}^{K} M_{ik}$. Group centered variables will be denoted with a C superscript. For example, $M_{ik}^{C} = M_{ik} - M_{i\cdot}$.[2]

To fit a model that most closely aligned with the most general model presented in this paper, the model for the effect of instance on the mediator, is

[2] The X variable is not mean-centered because in this design every individual experiences the same number or proportion of replicates in each instance. In this case $X_{i\cdot}$ will be constant across i. However for designs where X is not constant across i centering should be considered (Enders & Tofighi, 2007).

$$M_{ik} = a_{0i} + (a_1 + a_2 W_i) X_{ik} + e_{M_{ik}}$$
$$e_{M_{ik}} \sim N(0, \sigma^2_M)$$
$$a_{0i} = \alpha_{00} + \alpha_{01} W_i + u_{a_{0i}}$$
$$u_{a_{0i}} \sim N(0, \tau^2_{a_0})$$

For individual i at replicate k, M_{ik}, is a linear function of a person specific intercept a_{0i}, the instance for person i in replicate k, X_{ik}, and a person and replicate specific error term $e_{M_{ik}}$. The effect of X_{ik} is linearly moderated by W_i. The regression weight for the instance, $a_1 + a_2 W_i$, represents the conditional effect of instance on the mediator. The errors in this model are assumed to be independent and identically distributed with a mean of zero and a common variance σ^2_M. In the Level-2 model for the individual specific intercept, each individual's intercept a_{0i} is the combination of the population mean α_{00}, the moderator W_i with weight α_{01}, and an individual deviation from that predicted intercept $u_{a_{0i}}$. This deviation is assumed to have a normal distribution with mean 0 and variance $\tau^2_{a_0}$

The model for the simultaneous relationship between M and X and the outcome Y is

$$Y_{ik} = c'_{0i} + (c'_1 + c'_2 W_i) X_{ik} + (b_1 + b_2 W_i) M^C_{ik} + e_{Y'_{ik}} \tag{33}$$
$$e_{Y'_{ik}} \sim N(0, \sigma^2_{Y'})$$
$$c'_{0i} = \gamma'_{00} + \beta_{01} M_{i\cdot} + \beta_{02} W_i + u_{c'_{0i}}$$
$$u_{c'_{0i}} \sim N(0, \tau^2_{c'_0}).$$

The outcome for individual i in replicate k is predicted by the instance X_{ik} and the mediator M_{ik} where each of these effects is linearly moderated by W_i. Again, the intercept is individual specific, which allows responses from the same individual to be more similar to each other, but fixes the relationships between Y, X, and M to be the same across all individuals with the same value of W. The coefficient for instance, $c'_1 + c'_2 W_i$ represents the conditional direct effect (i.e. the effect of instance on the outcome holding the mediator constant and the moderator at a specific value). The coefficient for the mediator $b_1 + b_2W_i$ represents the relationship between the mediator and the outcome holding instance constant and at a specific value of the moderator. Similar to the model for M_{ik}, the model for the individual specific intercept includes a population mean γ'_{00}, a weight for the average level of the mediator β_{01}, a weight for the moderator β_{02}, and an individual deviation from that predicted value $u_{c'_{0i}}$ with variance $\tau^2_{c'_0}$.

The conditional direct effect is $c'_1 + c'_2W$, and the conditional indirect effect is $(a_1 + a_2W_i)(b_1 + b_2W_i)$. A simpler model could be fit where fewer paths are moderated. For a more complete description of estimation and inference in these models see Bauer, Preacher, and Gil (2006) and Rockwood and Hayes (2022). Zyphur, Zhang, Preacher, and Bird (2018) also discusses these models in a multilevel SEM framework.

Connection to Regression Based Method

Though the presentation of the multilevel models is different than the regression model, in the special case of a two-instance repeated-measure design, the regression model and the multilevel model can be equivalent and provide the same point estimates for each parameter and very similar standard error estimates.

To see the similarities between the two models, the equations for the multilevel model can be expressed as a single equation for each outcome, rather than broken into the Level 1 and Level 2 components.

$$M_{ik} = \alpha_{00} + \alpha_{01}W_i + u_{a_{0i}} + (a_1 + a_2W_i)X_{ik} + e_{M_{ik}}$$

Plugging in $X_{ik} = 0$ and $X_{ik} = 1$ for each equation shows the equivalence.

$$(M_{ik}|X_{ik} = 0) = \alpha_{00} + \alpha_{01}W_i + u_{a_{0i}} + e_{M_{i1}}$$

$$(M_{ik}|X_{ik} = 1) = (\alpha_{00} + a_1) + (\alpha_{01} + a_2)W_i + u_{a_{0i}} + e_{M_{i2}}$$

Compared to Equations 13 and 14 these two sets of equations are equivalent such that:

$$a_{10} = \alpha_{00}$$

$$\begin{aligned}
a_{20} &= \alpha_{00} + a_1 \\
a_{11} &= \alpha_{01} \\
a_{21} &= \alpha_{01} + a_2 \\
e_{M_{1i}} &= u_{a_{0i}} + e_{M_{ik}} \\
e_{M_{2i}} &= u_{a_{0i}} + e_{M_{ik}}
\end{aligned}$$

The parameters which are used to estimate the conditional indirect effect are the same in each type of model. The effect of instance on the mediator when the moderator is zero is $a_{10} - a_{20} = a_0$, and the degree to which W moderates the effect of instance on the mediator is $a_{11} - a_{21} = a_1$.

A similar process can be followed with the model of the outcomes; however the model from Equations 18 and 19 are only equivalent to the multilevel model under specific assumptions. The multilevel model equations are

$$\begin{aligned}
Y_{ik} &= \gamma'_{00} + \beta_{01} M_{i\cdot} + \beta_{02} W_i + u_{c'_{0i}} + (c'_1 + c'_2 W_i) X_{ik} \\
&\quad + (b_1 + b_2 W_i) M^C_{ik} + e_{Y_{ik}} \\
(Y_{ik} | X_{ik} = 0) &= \gamma'_{00} + \beta_{01} M_{i\cdot} + \beta_{02} W_i + u_{c'_{0i}} + (b_1 + b_2 W_i) M^C_{ik} + e_{Y_{i1}} \\
(Y_{ik} | X_{ik} = 1) &= (\gamma'_{00} + c'_1) + \beta_{01} M_{i\cdot} + (\beta_{02} + c'_2) W_i + u_{c'_{0i}} + (b_1 + b_2 W_i) M^C_{ik} + e_{Y_{i2}}
\end{aligned}$$

and the regression based equations are

$$Y_{1i} = c'^{*}_1 + (b_{11} + b_{13} W_i) M_{1i} + b_{12} W_i + e_{Y_{1i}}$$

$$Y_{2i} = c'^{*}_2 + (b_{21} + b_{23} W_i) M_{2i} + b_{22} W_i + e_{Y_{2i}}$$

The multilevel equations involve a term $\beta_{01} M_{i\cdot}$ whereas the regression based equations do not. This is because the regression based equations are used for estimating within-person (or Level 1) indirect effects, and because β_{01} is used for Level 2 indirect effects, it is not needed. Next, the regression based equations allow for the interaction between instance and the mediator, allowing $b_{11} \neq b_{21}$ and $b_{13} \neq b_{23}$. If it is assumed that

$b_{11} = b_{21}$ and $b_{13} = b_{23}$ then the following equivalences hold:

$$\begin{aligned} c_1'^{*} &= \gamma_{00}' \\ c_2'^{*} &= \gamma_{00}' + c_1' \\ b_{11} &= b_{21} = b_1 \\ b_{13} &= b_{23} = b_2 \\ e_{Y_{1i}'} &= u_{c_{0i}'} + e_{M_{ik}} \\ e_{Y_{2i}'} &= u_{c_{0i}'} + e_{M_{ik}} \end{aligned}$$

This shows under what assumptions the regression based approach and multilevel approach are equivalent. The multilevel model presented assumes that there is no interaction between instance and the mediator in the model for Y. This interaction could be included in the model; however, no literature on conditional process analysis in multilevel modeling does so. The regression based approach can be adjusted to accommodate this assumption, and MEMORE can estimate this model by adding the argument `xmint = 0`.

Though the methods proposed in this paper can be expressed as equivalent to multilevel approaches, there are still three distinct advantages of the the proposed methods. The first is ease of use: MEMORE is a very easy to use and intuitive software which can be used by researchers across a variety of fields without in-depth knowledge of MLM software. Software for moderated mediation in multilevel models is limited in availability (e.g., MLMED, (Rockwood, 2018)), and requires complex data restructuring prior to analysis (Bauer et al., 2006). In addition, without tools like MEMORE researchers have to identify which coefficients to combine and test, leading to potential errors. MEMORE automates the estimation of complex functions of parameters, providing the desired statistical output directly without additional, potentially error prone, computational steps. A second advantage of the proposed approach over multilevel models is the use of an OLS estimator compared to a maximum likelihood estimator. Maximum likelihood estimation relies on stronger assumptions and requires larger sample sizes (Maxwell & Delaney, 2004), so using OLS estimation may be preferred for the types of data often available in psychological research. The final benefit of the proposed method is the availability of bootstrapping as an inferential method for the index of moderated mediation and conditional indirect effects. In multilevel mediation it is common to use Monte Carlo confidence intervals (also available in MEMORE with the `mc = 1` option); however, this method relies on the assumption that the coefficients are normally distributed with constant variance (Selig & Preacher, 2009). Bootstrapping does not rely on the same assumptions and has shown been shown to perform as well as or better than Monte Carlo confidence intervals in repeated-measures designs (Montoya, 2022; Yzerbyt et al., 2018). Research in the context of between-subject designs has shown that Monte Carlo confidence intervals under-perform with non-normal distributions (Biesanz et al., 2010). No existing software conducts multilevel moderated mediation with bootstrap confidence intervals, as methods for bootstrapping multilevel data are complex to implement (Sharples & Page-Gould, 2016).

Extensions using MLM approaches

Multilevel modeling is a more general approach to conditional process analysis than the specific approach described in this paper which uses difference scores and linear regression. The two approaches will produce the same estimates under certain assumptions, but multilevel models can be generalized far beyond the two-instance repeated-measures case. Multilevel models would be able to accommodate multiple replicates for each individual in each instance. They can accommodate multiple conditions, or continuous, rather than dichotomous, X, including time as X. The moderator could be Level 1 or Level 2. All models in this section had fixed slopes and random intercepts, but multilevel models can also have random slopes. This adds additional complexity to the model which can be useful when theory suggests that the slopes may vary by individual. Including random slopes also lends better to the interpretation of the model as one of intra-individual change, since the change is measured at the individual level using the individual random coefficients. Literature which discusses more general applications of conditional process models with multilevel models include Bauer et al. (2006), Rockwood and Hayes (2022), Zyphur et al. (2018). The MLMED macro can by used to fit these models easily in SPSS (Rockwood, 2018; Rockwood & Hayes, 2022).

Discussion

Previous research on conditional process analysis has primarily focused on between-subjects designs; however, repeated-measures designs are very common in psychology and other social-behavioral sciences. This paper has integrated the research on mediation and moderation analysis to describe how to conduct conditional process analysis in two-instance repeated-measures designs. MEMORE, an easy-to-use macro for SPSS and SAS, was introduced as a way to fit these models, conduct inference and probe conditional effects. This development allows researchers who are investigating questions of conditional processes to use two-instance repeated-measures designs.

In this section, I discuss two broad limitations of the current approach and how these can be extended in future research. First, there is great need to more clearly understanding causal inference for conditional process analysis in two-instance repeated-measures designs. Second, there are limitations to the model of change due to both design and analytic choices which can be expanded using more complex designs and extensions of existing SEM approaches. Finally, I discuss potential extensions of the current proposed model which would continue to improve the alignment of available methods for between-subject and repeated-measures designs.

Causal Inference

As noted throughout this paper, there is a growing literature on causal inference for mediation in both between-subject and repeated-measures designs. Prior work on two-instance repeated-measures designs for mediation exist (Josephy et al., 2015) and also for moderated mediation in between-subjects designs (Qin & Wang, 2023), but their combination has not yet been discussed. This would help to formalize the assumptions needed for causal inference about the index of moderated mediation and conditional indirect effects in these designs. Additionally, once a clear set of assumptions is established, there is great need to develop sensitivity analyses to evaluate violations of such assumptions, similar to ? (?) but in repeated-measures designs.

A topic of much discussion in the mediation literature (whether between-subjects or repeated-measures) is concern over "cross-sectional mediation" where some or all measures (X, M, and Y) are measured "at the same time." Temporality of measurement is important for causal inference; however, it is difficult to be overly prescriptive about how much time must pass for events to be considered to occur at different "times." Many researchers criticize mediation analyses which use measures of the mediator and outcome taken in a single sitting (Bullock & Green, 2021; Maxwell & Cole, 2007); yet, other research focuses on how responding to one measure affects responses to another within the same sitting (e.g., order effects in measurement and priming; Strack, 1992; Şahin, 2021). Processes underlying certain psychological phenomena can happen in an instant or over prolonged periods of time. Greater rationale for the timing of observations should be given by researchers conducting mediation analysis, and researchers may seek to expand their research designs to measure variables at many time points, allowing for a clearer tracking of trajectories over time.

Even in cross-sectional contexts or pre-post test contexts, there are approaches to improving the evidence for temporal precedence (O'Rourke & MacKinnon, 2018). For example, questionnaire stems can ask participants to recall specific time periods (e.g., today, over the last week, over the last month). If mediator and outcome measures must be taken in the same session, researchers should consider aligning their question stems with the hypothesized time course of the effects of interest. This can be improved through using methods which track events specifically (e.g., Timeline Followback).

Models of Change

Difference scores have long been criticized as unreliable measures of growth (Lord, 1956) because they contain measurement error and are insensitive to artifacts such as regression to the mean, ceiling effects, and floor effects (Bonate, 2000; Cronbach & Furby, 1970; Lord, 1963; Twisk & Proper, 2004). However, many have argued that the merits of difference scores are greater than some researchers give them credit for (Rogosa, 1995; Thomas & Zumbo, 2012; Zumbo, 1999). Concerns about these types of artifacts do apply to the analysis described in this paper; however, one method for addressing this concern would be to include baseline measures (e.g., M_1 or Y_1) as moderators. This approach would allow for the estimation of autoregressive effects, rather than the assumption that they are perfectly stable. Alternative approaches, such as the ANCOVA model proposed by Valente and MacKinnon (2017) can be used for pre-post designs, but not for AB/BA cross-over designs.

By only measuring individuals twice, the model of change is fixed to be a linear model regardless of the analytical technique used. There may be many cases where change is expected to occur non-linearly, and would thus require more time points. For example, if a researcher were interested in understanding the effect of a temporary pain reliever (e.g., ibuprofen). Each participant could come into the lab and received a different dose of the drug ranging from 0mg to 800mg in 100mg increments. This would mean that each participant is measured 9 times. Given 9 repeated measurements there are a variety of models of change that could be fit to an individual's data: for example, a curvilinear relationship such that the decrease in pain with each 100mg increase slowly flattens out. This type of the theory could not be examined with the two-instance repeated-measures design, because individuals are only measured twice. This means that the difference between the two instances is inherently treated as linear change, and in order to examine more complex models of change a more complex design and analytic approach are required.

There are a variety of SEM methods which have been used to describe mediation processes with repeated-measures data (e.g., latent difference and latent growth curve models). With the exception of multilevel SEM, these methods have yet to be generalized to conditional process models. SEMs have many advantages over regression models: they can include measurement models to account for measurement error, estimate covariance between parameters estimated in different outcome models, and provide global model fit indices. Additionally, the SEM approaches can handle longitudinal change more flexibly than regression models. Generalizing mediation models for latent difference score (MacKinnon, 2008; Preacher, 2015; Selig & Preacher, 2009; Wu, Selig, & Little, 2013) and latent growth curve models (Cheong, MacKinnon, & Khoo, 2003; Cheong, 2011; Cole & Maxwell, 2003; Selig & Preacher, 2009) to conditional process models would be advantageous as researchers could utilize the benefits of SEMs and conditional process models for repeated-measures data all at the same time.

Model Extensions

While the proposed model in this paper takes a big step in increasing the range of conditional process models available to researchers for two-instance repeated-measures designs, there are still more models available for between-subject designs. One clear potential extension includes multiple moderators in the conditional process model for two-instance repeated-measures designs. As described in Hayes (2018a) and Montoya (2019), multiple moderation can occur in two ways: multiplicative moderation, where the moderators interact with each other, and additive moderation, where the moderators do not interact. Additive moderation is a simpler case, and if there were to be additional moderators in the conditional process models described in this paper, the conditional effects could just be linear functions of the moderators. If the moderators interact with each other, then this interaction term would need to be included in the conditional effects. This could lead to very complex functions defining the conditional indirect effect, especially if multiple paths in the mediation model are moderated by multiple moderators. Hayes (2018a) describes a variety of conditional process models with multiple moderators for between-subjects designs. All of these models could be estimated in the two-instance repeated-measures design using the principles described in this paper. The indices of partial, conditional, and moderated moderated mediation (Hayes, 2018b) could be generalized to two-instance repeated-measures designs.

Summary

Based on the methods proposed in this paper researchers can now conduct conditional process analysis in two-instance repeated-measures designs. I overviewed the basics of conditional process analysis by describing it's implementation in between-subjects designs, an area many researchers are more familiar with. I described the recent developments in mediation and moderation for two-instance repeated-measures designs. I then described how to combine the mediation and moderation approaches to estimate conditional process models. Additionally, I defined a variety of statistics of interest: the index of moderated mediation and conditional indirect effects. These statistics can be used to make inference about the population for effects specifically of interest in moderated mediation hypotheses. Next I provided an applied example of implementing these analyses in psychological data using the MEMORE macro for SPSS and SAS. This example (and a second example in Appendix A) showcase the cases in which MEMORE can be used and a variety of questions that can be answered using this analysis technique. In addition, they demonstrate the limitations of causal inference depending on design and analysis. I described a multilevel model that could be used for estimating conditional indirect effects in two-instance repeated-measures design. Finally, I discussed the limitations and future direction of the analysis technique and the two-instance repeated-measures design, specifically focusing on causal inference and models of change. This paper opens the door for a variety of extensions of this model and for there to be increased discussion about the connections between analytical approaches for between-subjects and repeated-measures designs, especially within the context of conditional process analysis.

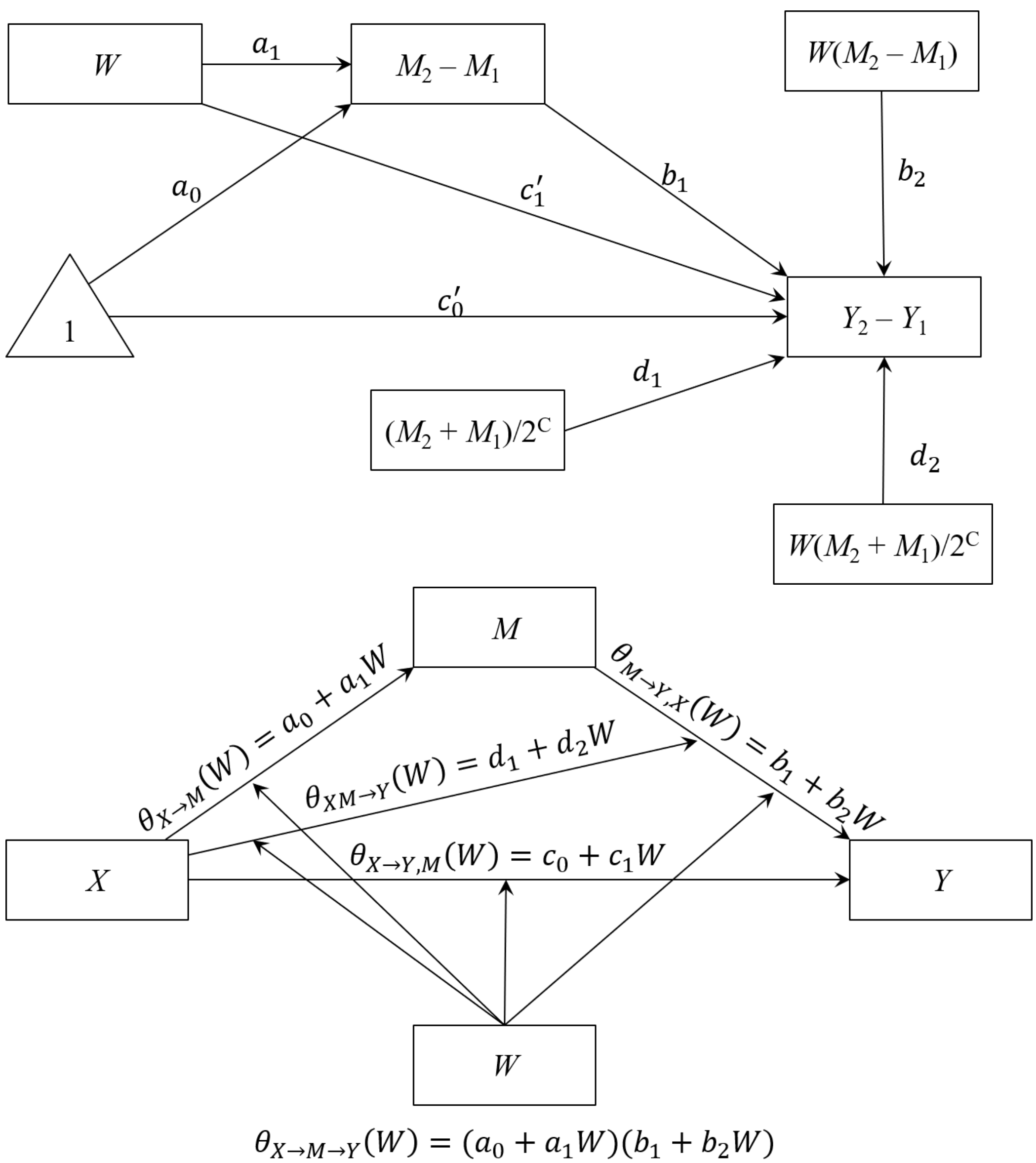

Figure 2. Statistical (top) and Conceptual (bottom) path diagram for the general conditional process model, where W moderates all paths in the mediation model

Table 2

The Index of Moderated Mediation and Conditional Indirect Effect with Second Order Asymptotic Variances for Three Moderated Mediation Models

Model	Estimating	Index of Moderated Mediation		Conditional Indirect Effect	
	Equations	Estimate	Asymptotic Variance	Estimate	Asymptotic Variance
First Stage	15, 27/11	a_1b_1	$a_1^2 s^2_{\hat{b}_1} + b_1^2 s^2_{\hat{a}_1} + s^2_{\hat{a}_1} s^2_{\hat{b}_1}$	$(a_0 + a_1W)b_1$	$(a_0 + a_1W)^2 s^2_{\hat{b}_1} + (b_1^2 + s^2_{\hat{b}_1})(s^2_{\hat{a}_0} + 2s_{\hat{a}_0,\hat{a}_1}W + s^2_{\hat{a}_1}W^2)$
Second Stage	7, 30/21	a_0b_2	$a_0^2 s^2_{\hat{b}_2} + \hat{b}_2^2 s^2_{\hat{a}_0} + s^2_{\hat{a}_0} s^2_{\hat{b}_2}$	$a_0(b_1 + b_2W)$	$(b_1 + b_2W)^2 s^2_{\hat{a}_0} + (a_0^2 + s^2_{\hat{a}_0})(s^2_{\hat{b}_1} + 2s_{\hat{b}_1,\hat{b}_2}W + s^2_{\hat{b}_2}W^2)$
General (First and Second) Stage	15, 21	N/A	N/A	$(a_0 + a_1W)(b_1 + b_2W)$	$(a_0 + a_1W)^2(s^2_{\hat{b}_1} + 2s_{\hat{b}_1,\hat{b}_2}W + s^2_{\hat{b}_2}W^2)+$ $(b_1 + b_2W)^2(s^2_{\hat{a}_0} + 2s_{\hat{a}_0,\hat{a}_1}W + s^2_{\hat{a}_1}W^2)+$ $(s^2_{\hat{b}_1} + 2s_{\hat{b}_1,\hat{b}_2}W + s^2_{\hat{b}_2}W^2)(s^2_{\hat{a}_0} + 2s_{\hat{a}_0,\hat{a}_1}W + s^2_{\hat{a}_1}W^2)$

$s^2_{\hat{b}_j}$ is the sampling variance of the coefficient b_j

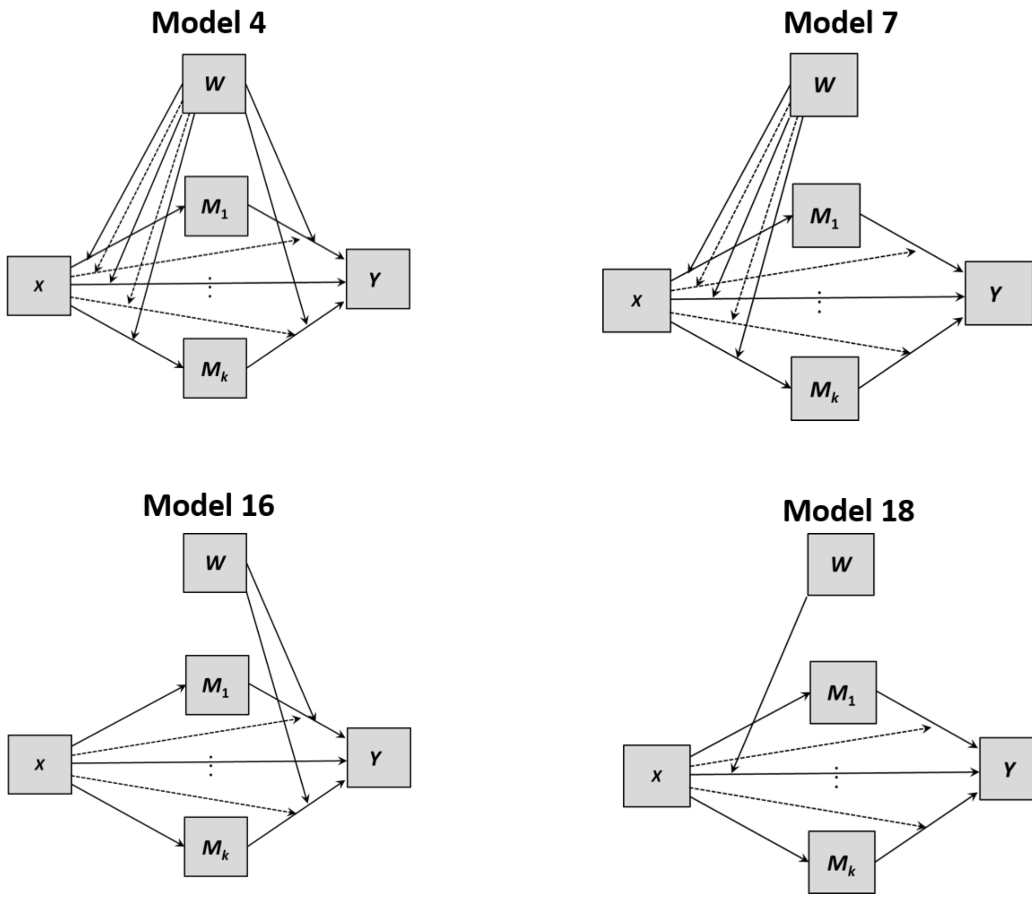

Figure 3. Examples of models in MEMORE. Model 4 allows moderation on all paths. Model 7 allows moderation on a-paths, d-paths, and the direct effect. Model 16 allows moderation on only the b-paths. Model 18 allows for moderation only on the direct effect.

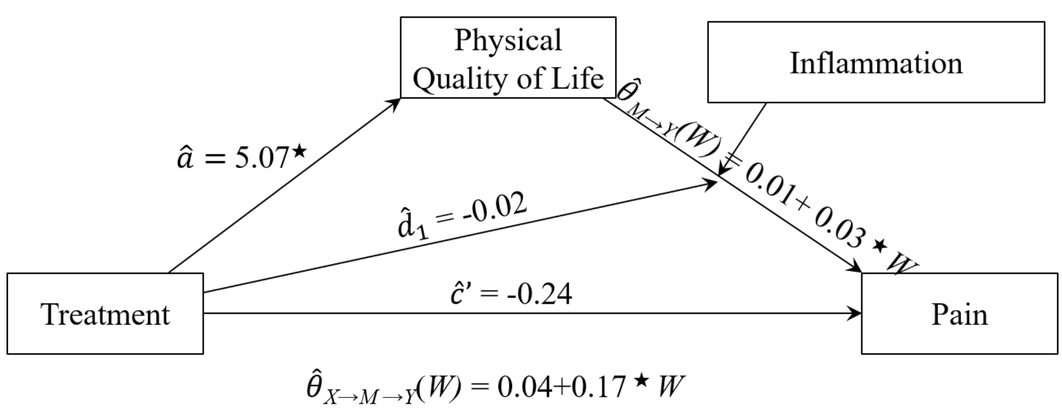

Figure 4. Diagram for Lasselin et al., (2016) example with a second-stage moderated mediation. * indicates statistically significant paths at $\alpha = .05$.

References

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendation. *Psychological Methods*, *11*(2), 142–163. doi: 10.1037/1082-989X.11.2.142

Bell, Z., Shader, T., Webster-Stratton, C., Reid, M. J., & Beauchaine, T. P. (2017). Improvements in negative parenting mediate changes in children's autonomic responding following a preschool intervention for adhd. *Clinical Psychological Science*, *6*(1), 1–11. doi: 10.1177/2167702617727559

Biesanz, J. C., Falk, C. F., & Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects. *Multivariate Behavioral Research*, *45*, 661–701. doi: 10.1080/00273171.2010.498292

Bonate, P. L. (2000). *Analysis of pretest-posttest designs.* New York: Chaptman & Hall / CRC.

Bullock, J. G., & Green, G. P. (2021). The failings of conventional mediation analysis and a design-based alternative. *Advances in Methods and Practices in Psychological Science*, *4*(4), 1–18. doi: 10.1177/25152459211047227

Cheong, J. (2011). Investigation of mediational processes using latent growth curve modeling. *Structural Equation Modeling*, *18*, 195-211.

Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using latent growth curve modeling. *Structural Equation Modeling*, *10*(2), 238. doi: $10.1207/S15328007SEM1002_5$

Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computer & Education*, *57*(2), 1825–1835. doi: 10.1016/j.compedu.2011.02.004

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), 558–577. doi: 10.1111/acer.12126

Coutts, J. (2023). *Enhancing the specification, testing, and interpretation of conditional indirect effects* (Dissertation). The Ohio State University.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change" — or should we? *Psychological Bulletin*, *74*(1), 68–80. doi: 10.1037/h0029382

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

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. doi: 10.1037/1082-989X.12.2.121

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science*, *10*(2), 87–99. doi: 10.1007/s11121-008-0109-6

Fossum, J. L. (2023). *Impacts of model specification on statistical power and type i error rate in moderated mediation analysis* (Unpublished doctoral dissertation). University of California, Los Angeles.

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, *50*(1), 1 – 22. doi: 10.1080/00273171.2014.962683

Hayes, A. F. (2018a). *Introduction to mediation, moderation, and conditional process analysis* (2nd ed.). New York, NY: Guilford Press.

Hayes, A. F. (2018b). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*, 4–40. doi: 10.1080/03637751.2017.1352100

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

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. doi: 10.1177/0956797613480187

Imai, K., Keele, K., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. doi: 10.1214/10-STS321

Josephy, H., Vansteelandt, S., Vanderhasselt, M.-A., & Loeys, T. (2015). Within-subjects mediation analysis in AB/BA crossover designs. *International Journal of Biostatistics*, *11*(1), 1–22. doi: 10.1515/ijb-2014-0057

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subjects designs. *Psychological Methods*, *6*(2), 115-134.

Judd, C. M., McClelland, G. H., & Smith, E. R. (1996). Testing treatment by covariate interactions when treatment varies within subjects. *Psychological Methods*, *1*, 366 - 378.

Kenny, D. A., & Judd, C. M. (2014). Power anomalies in testing mediation. *Psychological Science*, *25*(2), 334 – 339.

Lasselin, J., Kemani, M. K., Kanstrup, M., Olsson, G. L., Axelsson, J., Andreasson, A., ... Wicksell, R. K. (2016). Low-grade inflammation may moderate the effect of behavioraltreatment for chronic pain in adults. *J Behav Med*, *39*, 916 – 924.

Loeys, T., Talloen, W., Goubrt, L., Moerkerke, B., & Vansteelandt, S. (2016). Assessing moderated mediation in linear models requires fewer confounding assumptions than assessing mediation. *British Journal of Mathematical and Statistical Psychology*, *69*(3), 352 – 374.

Lord, F. M. (1956). The measurement of growth. *Educational and Psychological Measurement*, *16*(4), 421 – 437. doi: 10.1177/001316445601600401

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison, Wisconsin: University of Wisconsin Press.

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

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*, 99 – 128.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. doi: 10.1037/1082-989X.12.1.23

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Montoya, A. K. (2018). *Conditional process analysis in two-instance repeated-measures designs* (Dissertation). The Ohio State University.

Montoya, A. K. (2019). Moderation analysis in two-instance repeated measures designs: Probing methods and multiple moderator models. *Behavior Research Methods*, *53*(1), 61–82. doi: 10.3758/s13428-018-1088-6

Montoya, A. K. (2022). Selecting a within- or between-subject design for mediation: Validity, causality, and statistical power. *Multivariate Behavior Research*, *44*(3), 616–636. doi: 10.1080/00273171.2022.2077287

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. doi: 10.1037/met0000086

Montoya, A. K., Hayes, A. F., & Gomez, N. (Under Revision). Mediation analysis in the two condition pretest-postest design: A treatment-as-moderator conditional process approach.
doi: 10.31234/osf.io/rfvs7

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, *89*(6), 852–863. doi: 10.1037/0022-3514.89.6.852

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*(2), 171–181. doi: 10.15288/jsad.2018.79.171

Pearl, J. (2001). *Direct and indirect effects.* San Francisco, CA: Morgan Kaufman.

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*, 825–852. doi: 10.1146/annurev-psych-010814-015258

Qin, X., & Wang, L. (2023). Causal moderated mediation analysis: Methods and software. *Behavior Research Methods*. doi: 10.3758/s13428-023-02095-4

Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent-child interaction coding system: Standardization and validation. *Journal of Consulting and Clinical Psychology*, *49*(2), 245–250. doi: 10.1037//0022-006x.49.2.245

Rockwood, N. J. (2018). *Advancing the formulation and testing of multilevel mediation and moderated mediation models* (Unpublished master's thesis). The Ohio State University, Columbus, OH.

Rockwood, N. J., & Hayes, A. F. (2022). Multilevel mediation analysis. In *Multilevel modeling methods with introductory and advanced applications* (A. A. O'Connell, D. B. McCoach, & B. A. Bell (Eds.) ed.). Charlotte, NC: Information Age Publishing.

Rogosa, D. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. M. Gottman (Ed.), *The analysis of change.* Mahwah, NJ: Lawrence Erlbaum Associates.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2–3), 144–164.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (1st ed.). Boston, MA: Houghton Mifflin Company.

Sharples, A., & Page-Gould, E. (2016). *Accurate indirect effects in multilevel mediation for repeated measures data.* Retrieved from `https://www.page-gould.com/r/indirectmlm/` (Talk presented at the Society for Personality and Social Psychology Annual Convention)

Stone, C. A., & Sobel, M. E. (1990). The robustness of estimates of total indirect effects in covariance structure models estimated by maximum likelihood. *Psychometrika*, *55*(2), 337–352. doi: 10.1007/BF02295291

Strack, F. (1992). "Order effects" in survey research: Activation and information functions of preceding questions. In N. Schwarz & S. Sudman (Eds.), *Context effects in social and psychological research* (pp. 23–34). New York, NY: Springer New York. Retrieved from `https://doi.org/10.1007/978-1-4612-2848-6`$_3$ doi: 10.1007/978-1-4612-2848-6$_3$

Thomas, D. R., & Zumbo, B. D. (2012). Difference scores from the point of view of reliability and repeated-measures anova: In defense of difference scores for data analysis. *Educational and Psychological Measurement*, *72*(1), 37–43. doi: 10.1177/0013164411409929

Twisk, J., & Proper, K. (2004). Evaluation of the results of a randomized controlled trial:How to define change between baseline and followup. *Journal of Clinical Epidemiology*, *57*, 223 – 228.

Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest-posttest control group design. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(3), 428–450.

Valente, M. J., Rijnhart, J. J. M., & Gonzalez, O. (2023). A novel approach to estimate moderated treatment effects and moderated mediated effects with continuous moderators. *Psychological Methods*. doi: 10.1037/met0000593

Vanderweele, T. J., & Vansteelandt, S. (2009). Conceptual issues concerning mediation, intervations, and composition. *Statistics and its Interface*, *2*, 457 – 468.

Wu, W., Selig, J. P., & Little, T. D. (2013). Longitudinal data analysis. In T. D. Little (Ed.), *Handbook of quantitative methods.* Oxford.

Yzerbyt, V. Y., Muller, D., Batailler, C., & Judd, C. M. (2018). New recommendations for testing indirect effects in mediational models: The need to report and test component paths. *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, *115*(6), 929–943. doi: 10.1037/pspa0000132

Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of experimental social psychology*, *40*, 424–431. doi: 10.1016/j.jesp.2003.10.001

Zumbo, B. D. (1999). The simple difference score as an inherently poor measure of change: Some reality, some mythology. In B. Thompson (Ed.), *Advances in social science methodology, 5.* Greeniley ich, CT: JAI Press.

Zyphur, M. J., Zhang, Z., Preacher, K. J., & Bird, L. J. (2018). Moderated mediation in multilevel structural equation models: Decomposing effects of race on math achievement within versus between high schools in the united states. In S. E. Humphrey & J. M. LeBreton (Eds.), *The handbook of multilevel theory, measurement, and analysis* (pp. 473–494). Washington, D. C.: American Psychological Association.

Şahin, M. D. (2021). Effect of item order on certain psychometric properties: A demonstration on a cyberloafing scale. *Frontiers in Psychology*, *12*. doi: 10.3389/fpsyg.2021.590545

Appendix A

Applied Example: Bell, Shader, Webster-Stratton, Reid, Beauchaine (2017)

Bell, Shader, Webster-Stratton, Reid, and Beauchaine (2017) investigated whether a family behavioral intervention impacts externalizing symptoms in children with ADHD, and whether this effect is mediated by parenting behaviors. The training was administered in 20 two-hour weekly sessions. In this study, families were randomly assigned to either receive the treatment starting immediately at Week 1 or a delayed treatment, starting at Week 10. Each family was measured before the intervention period (Week 0), and then again 20 weeks later, meaning the delayed treatment group received about half the amount of treatment that the immediate group received at the time of assessment. The mediator of interest in the study was positive parenting behaviors (Robinson & Eyberg, 1981). Increases in positive parenting were expected to be more pronounced in the immediate treatment group compared to the delayed treatment group (i.e., treatment moderates the effect of time on the mediator). The outcome variable was cardiac pre-ejection period (PEP) reactivity. Dysfunctional PEP reactivity manifests as lower scores and can put children at risk for ADHD. Higher PEP reactivity is considered to be more "typical." The intervention was meant to increase PEP reactivity over time. Positive parenting and PEP reactivity were measured before the intervention period and at the end of 20 weeks. In the original paper, the researchers estimated separate mediation models for the immediate treatment group and the delayed treatment group, finding a significant indirect effect of time on PEP reactivity through parenting in the immediate treatment group but not the delayed treatment group. The analysis presented here provides a direct test of the difference between those indirect effects.

An alternative analytical approach could treat the treatment condition (between-subject factor) as the focal predictor in the mediation analysis using the ANCOVA approach described by (Valente & MacKinnon, 2017). In this approach, treatment condition would predict PEP reactivity at Week 20 indirectly through Positive

Parenting at Week 20 while controlling for PEP reactivity and Positive Parenting at Week 0. This approach provides an estimate of the total, direct, and indirect effects of treatment on the outcome at Week 20, but this method does not provide clear information about the direction of change over time within each group. Montoya, Hayes, and Gomez (Under Revision) demonstrates that the indirect effect of the ANCOVA model is equivalent to the index of moderated mediation generated by MEMORE under specific assumptions. Both analytic approaches may be appropriate, and the distinction is primarily in whether the focus is on the indirect effect of the repeated-measures factor as moderated by the between-subject factor (MEMORE approach) as compared to the indirect effect of the between-subject factor on follow-up measures while controlling for baseline measures (ANCOVA).

In this first-stage moderated mediation analysis, time (Week 0 vs. Week 20) is considered the independent variable, positive parenting is the mediator, and PEP reactivity is the outcome. Experimental condition (treatment vs. delayed) moderates the relationship between time and positive parenting. Additionally, the direct effect is allowed to be moderated, because there may be other factors that lead PEP to change over time and these other factors may depend on treatment condition. This is Model 12 in MEMORE, where the path from X to M (first stage) and the direct effect are allowed to be moderated.

The MEMORE code to estimate this model in SPSS is:

```
MEMORE y = PostPEP PrePEP /m = PostPos PrePos /w = cond
/model = 12 /plot = 1/seed = 104874.
```

The `y` and `m` arguments are used to provide the names of the variables for the outcome and mediators, respectively. The order determines the order of subtraction, here it is $post - pre$, which results in positive scores for variables that increase over time and negative scores for variables and decrease over time. The `w` argument provides the name of the moderator variable. The `model` argument is set to 12, allowing the path from X to M and direct effect to be moderated (See documentation for all model descriptions). The `plot` argument is set to 1, requesting plotting information. Finally, `seed` is set for reproducibility. The MEMORE output is available in Appendix D. The complete estimated model is represented in Figure A1.

The first section of the output provides information about model specification, computed variables, sample size, and random seed. The next section provides information about the total effect model, which is a simple moderation model where condition is allowed to moderate the total effect of time on PEP reactivity, where the sample estimated equation is

$$\widehat{Y_{2i} - Y_{1i}} = 0.0170 - 0.0355W_i \tag{34}$$

In this data, treatment group is coded such that the delayed intervention group is 0 and the immediate intervention group is 1. The intercept estimate is $\hat{c}_0 = 0.0170$, meaning that among those in the delayed treatment condition ($W = 0$), PEP reactivity is estimated to increase by 0.0170 units, but this change is not significantly different from zero and should not be generalized to the population ($t(97) = 0.0238$, $p = 0.98$). The coefficient for condition, provides the estimated difference in change over time comparing the delayed condition and immediate treatment condition. So, $\hat{c}_1 = -0.0355$ means the immediate treatment condition is estimated to decrease in PEP over time 0.0355 units more, compared to the delayed condition. While PEP was hypothesized to increase more in the immediate condition than delayed, this effect is in the opposite direction, but not statistically significant ($t(97) = -0.0349$, $p = 0.97$). Below the Model Summary section is the conditional effect of time on PEP reactivity in each of the conditions. Note that when condition is 0, this is the same as the intercept in the previous model. But if condition had been coded differently, for example -0.5 and 0.5 then the intercept and this conditional effect would provide distinct information. When condition is 1 (immediate treatment), PEP is estimated to decrease over time by 0.0185 units, but this effect is not statistically significant ($t(97) = -0.0256$, $p = 0.98$). Notice that this example is a case where the total effect is not statistically significantly moderated, but this does not necessarily mean that the indirect effect is not moderated. A lack of evidence of moderation on the total effect should not be a barrier to the pursuit of hypotheses about moderated indirect or direct effects (Hayes, 2022; O'Rourke & MacKinnon, 2018). It possible that both the indirect and direct effects could be moderated in opposite directions, resulting in no moderation of the total effect, or that the test of moderation of the total effect could have lower power than the test of moderation of the indirect effect (Kenny & Judd, 2014).

The next section of the MEMORE output provides information about the effect of time on positive parenting ($X \rightarrow M$). Equation 15 estimates an intercept and a coefficient for the moderator (W, treatment group). For positive parenting, the estimated intercept is $\hat{a}_0 = 5.73$: for those in the delayed treatment group ($W = 0$), positive parenting increased by 5.73 units on average between baseline and follow-up, but this effect was not statistically significant ($t(97) = 1.74$, $p = .08$). The coefficient for W is estimated as $\hat{a}_1 = 10.64$, meaning that those in the immediate treatment group increased 10.64 units more on positive parenting on average than those in the delayed treatment group, and this effect was statistically significant ($t(97) = 2.28$, $p = .03$). This result can be interpreted as evidence that the change over time on positive parenting is greater in the immediate treatment group than the delayed treatment group at the population level. The effect of X on M is moderated by W.

Next, Equation 16 estimates the effect of positive parenting on PEP reactivity. Additionally, the direct effect is allowed to be moderated by treatment. The estimated model in the data is

$$\widehat{Y_{2i} - Y_{1i}} = -0.28 - 0.48W_i + 0.05(M_{12i} - M_{11i}) - 0.01(\frac{M_{12i} + M_{11i}}{2})^C \qquad (35)$$

In this equation, the intercept is the predicted difference in PEP reactivity for a child with parents that showed no difference over time on positive parenting, are average on positive parenting, and are in the delayed treatment condition. The intercept estimate is $\hat{c'}_0 = -0.28$, and this effect is not statistically significantly different from zero

($t(95) = -0.38, p = .71$). The coefficient for the moderator estimates how much the direct effect differs for individuals in the immediate treatment condition compared to delayed, conditional on being average on positive parenting. The estimate is $\hat{c'}_1 = -0.48$, and it is not statistically significant ($t(95) = -0.46$, $p = .65$), so there is not sufficient evidence to suggest the direct effect is moderated.

The slope for positive parenting, $\hat{b}_1 = .04$, gives the relationship between positive parenting and PEP reactivity averaged over baseline and follow-up within the sample. This coefficient is statistically significant ($t(95) = 2.06$, $p = .04$) suggesting that parents who provide more positive parenting tend to have children with higher PEP reactivity at the population level. The coefficient for the average of the mediator reflects the difference in the relationship between the mediator and the outcome at baseline compared to follow-up. $\hat{d}_1 = -.01$, can be interpreted to mean that the estimated relationship between positive parenting and PEP reactivity is .01 units lower at follow-up compared to baseline, but this difference is not statistically significant ($t(95) = -0.2295$, $p = 0.82$): the relationship between positive parenting and PEP reactivity is not statistically significantly different at baseline compared to follow-up.

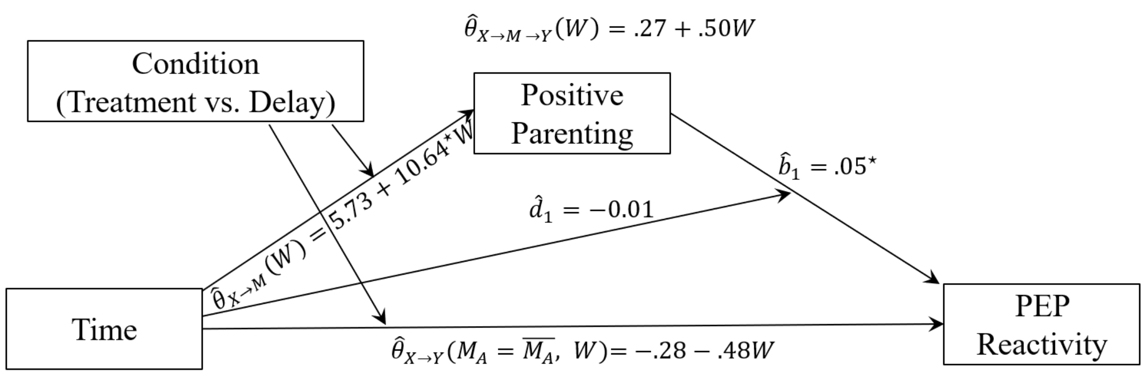

Figure A1. Diagram for Bell et al. (2018) First-Stage Conditional Process Model.* indicates statistically significant paths at $\alpha = .05$.

The product of the conditional effect of X on the mediator M ($\theta_{X \to M}(W)$) and the effect of M on Y ($\theta_{M \to Y}$), is the conditional indirect effect: $(a_0 + a_1 W)b_1$. The estimated conditional indirect effect through positive parenting is $(5.73 + 10.63W).05 = .27 + .50W$. MEMORE calculates the conditional indirect effects evaluated at each of the two values of the moderator (0 [delayed treatment] and 1 [immediate treatment]), as well as percentile bootstrap confidence intervals for these effects. See Table A1 for estimates of the conditional indirect and direct effects in each of the two conditions. For individuals in the

Table A1

Conditional Effects of Time on PEP Reactivity with 95% confidence intervals

Moderator Value	Indirect Effect			Direct Effect		
0 (Delayed Intervention)	0.2680	[-0.0062	0.6700]	-0.2510	[-1.6785	1.1764]
1 (Immediate Treatment)	0.7657	[0.0401	1.6979]	-0.7842	[-2.3834	0.8150]

Percentile bootstrap confidence intervals are used for the indirect effects. Parametric confidence intervals are used for the direct effects.

delayed onset condition, the indirect effect of time on PEP reactivity through positive parenting was not statistically significant. However, for those in the immediate treatment condition there was a significant indirect effect through positive parenting, where PEP reactivity increased over time through increases in positive parenting. Both conditional direct effects were not significantly different from zero, nor was there strong evidence that the direct effect is moderated by condition.

The hypothesized direction of change was supported by the data. Those in the immediate treatment condition experience positive indirect effects on PEP reactivity through positive parenting and this indirect effect was not observed among the delayed treatment group. To test the differences between these indirect effects, the index of moderated mediation and a bootstrap confidence interval for the index can be used. The index of moderated mediation in this case is $a_1 b_1$.

The index is estimated to be 0.4977 with a bootstrap confidence interval of -0.0243 to

1.3096. This means that the indirect effect through positive parenting is estimated to be 0.4977 units higher in the immediate treatment condition compared to the delayed treatment condition, based on this sample. But because the confidence interval includes zero, there is not sufficient evidence that these two indirect effects are significantly different from each other. So while one indirect effect was statistically significant, and the other was not, they are not statistically significantly different from each other, and there is not sufficient evidence to claim the indirect effect is moderated by treatment condition.

Causal Inference

There are important limitations to causal inference in this design and analysis. The linear models used to estimate effects must be correctly specified for valid causal inference. Temporal precedence in this specific example means that condition should precede the change in positive parenting, which it does because condition is assigned at baseline. Additionally, change in positive parenting should precede change in PEP reactivity. The measure of positive parenting (a 7 min free-play task) precedes the measure of PEP reactivity in the lab setting. So this assumption is also supported.

Participants are randomly assigned to condition, which aids causal inference for some parts of the model, but random assignment to condition does not ensure that all paths in the moderated mediation model are identified. By allowing condition as a moderator of both the path from instance (in this case time) to positive parenting and PEP reactivity, the moderation coefficients can be assumed to be unbiased causal estimates of the effect of condition on these effects (Loeys et al., 2016). Specifically, because the moderator is manipulated, the difference between conditional effects is identified, but the conditional effects themselves may still be biased. Because the index of moderated mediation relies on the moderation parameter (a_1) this helps support an argument that the index itself may be identified. However, the second parameter (b_1) also needs to be identified. The random assignment does not aid in the identification of b_1. The previously discussed assumptions from two-instance mediation should apply in this context (Josephy et al., 2015), specifically that the assumption that there are no confounders of the effect of positive parenting on PEP reactivity (A.3 and A.4) is needed. Similar to Josephy et al. (2015), if unmeasured upper-level confounders of the $M - Y$ relationship have an additive effect and there is no unmeasured heterogeneity in $X - M$, $M - Y$, and $X - Y$ effects, then upper-level confounders can be ignored in this case. Given this study design and the random assignment to condition, the only assumption required for causal inference for the index of moderated mediation is that there are no lower-level confounders of the effect of positive parenting and PEP reactivity. Whether these assumptions are reasonable is difficult to know without subject specific knowledge. Future research must develop sensitivity analyses for conditional process analysis in repeated-measures designs. Until then researchers must be exceedingly cautious about making causal claims based on their analyses.

While the assumptions required for causal inference for the index of moderated mediation are relatively few, the same is not true for the conditional indirect effects. By a similar logic to Loeys et al. (2016), if the index of moderated mediation is non-zero, so too are at least one of the conditional indirect effects, but beyond this, there is little hope of ensuring identification of these effects in this type of design. Specifically, because the model relies on a pre-post design, the basic assumption is that the only reason for change over time is the treatment itself (A1). For example, the coefficient a_0 provides an estimate of how much positive parenting changed over time, but does not account for how much of this change is due to the treatment vs. how much is due to other factors. So long as these other factors operate equally for those in the immediate and delayed treatments, this does not impact the estimate of the effect of timing of treatment (immediate vs. delayed), but does limit the claims about the conditional indirect effects as causal effects of treatment, and they should be interpreted with caution. This is emphasized by the X variable in the model being described as "Time" meaning that the effects of these variables in these models typically only reflect change over time. So the coefficient a_0 (and thus also $a_0 + a_1$) cannot be interpreted as causal effects in these designs, unless assuming there are no other factors that lead to change in positive parenting over time (A1). Because the conditional indirect effects also involve the b_1 path, the same no-confounding assumption described above would also be required (A3, A4). Overall, this suggests there are few conditions under which the conditional indirect effects themselves would be identified. However, this is not necessarily a limitation of the statistical method, so much as the difficulty of making causal inference from pre-post designs. The addition of random assignment to condition, however, aids in identification of the most important parameter of interest: the index of moderated mediation.

Appendix B

Fitting Conditional Process Analysis for Two-Instance Repeated-Measures Designs in R with lavaan

Using a structural equation modeling program for fitting the models described in this paper. The same linear regression equations can be fit simultaneously using maximum likelihood in a structural equation modeling program. In this appendix, the R package `lavaan` is used to fit the Applied Example in the manuscript (Lasselin et al., 2016) and the second applied example in Appendix A (Bell et al., 2017).

```
library(tidyverse)
```

```
Warning: package 'tidyverse' was built under R version 4.3.2
```

```
Warning: package 'lubridate' was built under R version 4.3.2
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.3     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.0
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```
library(lavaan)
```

```
Warning: package 'lavaan' was built under R version 4.3.2
```

```
This is lavaan 0.6-17
lavaan is FREE software! Please report any bugs.
```

```
library(foreign) #used for read.spss
```

```
Warning: package 'foreign' was built under R version 4.3.2
```

```
set.seed(1140909)
```

Lasselin et al. (2016)

```
      LasselinData <- read.spss("LasselinCPAData.sav", use.value.labels = FALSE,
                                                      to.data.frame = TRUE)
```

The above code reads in the data, assuming it is in the working directory.

First create variables which represent the difference in the mediator, difference in the outcome, and the mean-centered average of the mediators.

```
LasselinData <- LasselinData %>% transform(
      Ydiff = PostPain - PrePain,
      Mdiff = PostPCS - PrePCS,
      Mavg = (PostPCS + PrePCS)/2 - mean((PostPCS + PrePCS)/2)
)

LasselinData <- LasselinData %>% transform(
        MdiffW = Mdiff*inflame
)
```

The `lavaan` code to fit the model is below. This code uses Equations 7 and 17 (except omitted the moderated direct effect). The first line estimates a model of the difference in the mediators, and does not include the moderator (Eq. 7). The second line estimates a model of the difference in the outcomes, predicted by the difference between the mediators, the mean-centered average of the mediators, and the interaction between the moderator and the difference between the mediators (Eq. 17 dropping coefficient for W). This model allows the effect of the mediator on the outcome to be conditional on the moderator, but direct effect and the difference between the effect of M on Y in each instance are not allowed to be conditional on the moderator. The parameters are all labeled to align with the equations from the original paper (e.g., `a0`

$= a_0$). By naming the parameter estimates, they can be used to calculate new parameter estimates. For example the conditional indirect effect when W is one SD below the mean is called `indirectLow`, and when W is 1 SD above the mean it is called `indirectHigh`. These are calculated using the equations from the main manuscript (Table 2). The mean and standard deviation have to calculated in advance and typed in as part of the `lavaan` syntax, which is potentially error prone. These values can also be used to calculate the conditional effect of Mdiff on Ydiff.

```
mean(LasselinData$inflame)
```

```
[1] 0.0006282056
```

```
sd(LasselinData$inflame)
```

```
[1] 1.019242
```

The index of moderated mediation is also calculated and named `IMM`. The results match the manuscript.

```
LasselinModel <- '
Mdiff ~ a0*1
Ydiff ~ cp0*1+ b1*Mdiff +b2*MdiffW+ d1*Mavg

# Conditional Effects of Mdiff on Ydiff
MYLow := (b1+b2*(0.0006282056-1.019242))
MYMean :=(b1+b2*(0.0006282056))
MYHigh :=(b1+b2*(0.0006282056+1.019242))

# Conditional Indirect Effects
indirectLow := (a0)*(b1+b2*(0.0006282056-1.019242))
indirectMean :=(a0)*(b1+b2*(0.0006282056))
indirectHigh :=(a0)*(b1+b2*(0.0006282056+1.019242))

IMM := a0*b2
'
```

Next we can fit this model in `lavaan`

```
LasselinFit <- sem(model = LasselinModel, data = LasselinData, se = "bootstrap",
bootstrap = 5000)
```

```
Warning in lav_partable_vnames(FLAT, "ov.x", warn = TRUE): lavaan WARNING:
    model syntax contains variance/covariance/intercept formulas
    involving (an) exogenous variable(s): [Mdiff]; These variables
    will now be treated as random introducing additional free
    parameters. If you wish to treat those variables as fixed, remove
    these formulas from the model syntax. Otherwise, consider adding
    the fixed.x = FALSE option.
```

```
summary(LasselinFit, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 14 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                         7

  Number of observations                            78

Model Test User Model:

  Test statistic                                30.123
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.000

Parameter Estimates:

  Standard errors                            Bootstrap
  Number of requested bootstrap draws             5000
  Number of successful bootstrap draws            5000

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Ydiff ~
    Mdiff     (b1)    0.008    0.014    0.567    0.571   -0.018    0.038
    MdiffW    (b2)    0.033    0.011    3.073    0.002    0.012    0.055
    Mavg      (d1)   -0.021    0.012   -1.827    0.068   -0.046   -0.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    Mdiff     (a0)    5.066    0.947    5.351    0.000    3.245    6.912
   .Ydiff    (cp0)   -0.243    0.161   -1.507    0.132   -0.565    0.071
```

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Ydiff             1.335    0.183    7.292    0.000    0.940    1.667
    Mdiff            71.769   12.699    5.652    0.000   47.050   97.393

Defined Parameters:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    MYLow            -0.026    0.011   -2.309    0.021   -0.047   -0.002
    MYMean            0.008    0.014    0.568    0.570   -0.018    0.038
    MYHigh            0.042    0.023    1.864    0.062   -0.001    0.088
    indirectLow      -0.132    0.064   -2.060    0.039   -0.260   -0.007
    indirectMean      0.040    0.070    0.571    0.568   -0.096    0.187
    indirectHigh      0.213    0.114    1.869    0.062   -0.005    0.453
    IMM               0.169    0.059    2.880    0.004    0.058    0.286
```

The R package `lavaan` uses maximum likelihood, rather than ordinary least squares regression to estimate the model, which results in the same parameter estimates and slightly different standard error estimates. By default (i.e., without the `se = "bootstrap"` argument), `lavaan` calculates the standard error, z-value, and p-value for the conditional indirect effects using the delta method (Table 2), but this method is well known to be overly conservative. Therefore I recommend calculating percentile bootstrap confidence intervals for these parameters, which is achieved by adding the `se = "bootstrap"` option, and the number of bootstraps is specified in `bootstrap = 5000` (i.e., requesting 5,000 bootstraps, which is the same as the default in MEMORE. The `ci.lower` and `ci.upper` columns provide the lower and upper confidence limits of a percentile bootstrap confidence interval.

There are a number of options in this function which can be adjusted to align with options in `MEMORE`. For example, the level of the confidence interval can be changed using the `level` argument. The default is 0.95, but if a researcher wanted a 90% CI instead, they could add to the above command `level = 0.90` and this would adjust the level of the CI.

Just as in the Bell example, the total effect model was not estimated in the above `lavaan` model. A separate model is estimated below (Equivalent to Equation 33 in the main manuscript). The conditional effect of X on Y at the mean, one standard deviation below the mean, and one standard deviation above the mean of inflammation is also calculated.

```
LasselinTotalModel <- '
Ydiff ~ c0*1+c1*inflame

totallow := c0+c1*(0.0006282056-1.019242)
totalmean := c0+c1*(0.0006282056)
totalhigh := c0+c1*(0.0006282056+1.019242)
```

```r
'

LasselinTotalFit <- sem(LasselinTotalModel, data = LasselinData)
summary(LasselinTotalFit, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 1 iteration

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                         3

  Number of observations                            78

Model Test User Model:

  Test statistic                                 0.000
  Degrees of freedom                                 0

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Ydiff ~
    inflame   (c1)    0.402    0.130    3.097    0.002    0.147    0.656

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Ydiff     (c0)   -0.231    0.131   -1.759    0.079   -0.488    0.026

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Ydiff             1.345    0.215    6.245    0.000    0.923    1.768

Defined Parameters:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    totallow         -0.640    0.186   -3.435    0.001   -1.005   -0.275
    totalmean        -0.231    0.131   -1.757    0.079   -0.488    0.027
    totalhigh         0.179    0.186    0.958    0.338   -0.187    0.544
```

Compared to the results from MEMORE (using OLS) the parameter estimates are identical but the standard errors of the coefficients are slightly different due to the asympototic variance estimate used in maximum likelihood as compared to the finite sample variance estimate in OLS. The patterns of significance are the same.

In this case, the direct effect is not moderated the component approach is not accurate, and the conditional total effect must be calculated in a separate model. However, in the next example, both the indirect and direct effects are moderated, and the conditional total effect could be calculated in two ways: a separate model or by combining estimates in the moderated mediation.

This provides all of the relevant output for conducting a second-stage conditional process analysis without a moderated direct effect in R using `lavaan`. The basic principles of developing and programming these models can be used to generate similar code for any moderated mediation model fit in the MEMORE macro.

Example 1: Bell et al., 2017

```
BellData <- read.spss("BellCPAdata.sav", use.value.labels = FALSE, to.data.frame = TRUE)
```

The above code reads in the data, assuming it is in the working directory.

Using the difference score method used throughout the main manuscript, we first create variables which represent the difference in the mediator, difference in the outcome, and the mean-centered average of the mediators.

```
BellData <- BellData %>% transform(
      Ydiff = PostPEP - PrePEP,
      Mdiff = PostPos - PrePos,
      Mavg = (PostPos + PrePos)/2 - mean((PostPos + PrePos)/2)
)

MavgMean <- with(BellData, mean((PostPos + PrePos)/2))
```

The `lavaan` code to fit the model is below. This code uses Equations 16 and 17 from the main manuscript. The first line estimates a model of the difference in the mediators predicted by condition, allowing the effect of X on M to be conditional on the moderator (Eq. 16). The second line estimates a model of the difference in the outcomes, predicted by condition, the difference between the mediators, and the mean-centered average of the mediators (Eq. 17). This model allows the direct effect to be conditional on the moderator, but the effect of M on Y and the difference between the effect of M on Y in each instance are not conditional on the moderator.

The parameters are all labeled to align with the equations from the main manuscript (e.g., a0 $= a_0$). By naming the parameter estimates, they can be used to calculate new parameter estimates. For example the conditional indirect effect when $W = 0$ is called indirect0, and when $W = 1$ it is called indirect1. These are calculated using the equations from the main manuscript (Table 2). The index of moderated mediation is also calculated and named IMM.

```
BellModel <- '
Mdiff ~ a0*1+a1*cond
Ydiff ~ cp0*1+cp1*cond + b1*Mdiff + d1*Mavg
indirect0 := (a0+a1*0)*b1
indirect1 := (a0+a1*1)*b1
IMM := a1*b1
'
```

Next we can fit this model in lavaan

```
BellFit <- sem(model = BellModel, data = BellData, se = "bootstrap", bootstrap = 5000)
summary(BellFit, rsquare = TRUE, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 1 iteration

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                         8

  Number of observations                            99

Model Test User Model:

  Test statistic                                 8.183
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.004

Parameter Estimates:

  Standard errors                            Bootstrap
  Number of requested bootstrap draws             5000
  Number of successful bootstrap draws            5000

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Mdiff ~
```

```
    cond      (a1)   10.635    4.744    2.242    0.025    1.089   19.944
  Ydiff ~
    cond     (cp1)   -0.478    1.004   -0.476    0.634   -2.405    1.499
    Mdiff     (b1)    0.047    0.022    2.143    0.032    0.004    0.089
    Mavg      (d1)   -0.008    0.039   -0.209    0.834   -0.090    0.066

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff     (a0)    5.728    2.389    2.398    0.016    1.155   10.421
   .Ydiff    (cp0)   -0.278    0.676   -0.412    0.680   -1.636    0.998

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff           529.516  110.410    4.796    0.000  329.919  751.283
   .Ydiff            23.990    4.014    5.976    0.000   15.369   31.192

R-Square:
                   Estimate
    Mdiff             0.051
    Ydiff             0.047

Defined Parameters:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    indirect0         0.268    0.174    1.538    0.124   -0.001    0.668
    indirect1         0.766    0.426    1.796    0.073    0.058    1.722
    IMM               0.498    0.346    1.437    0.151   -0.020    1.299
```

The R package `lavaan` uses maximum likelihood, rather than ordinary least squares regression to estimate the model, which results in the same parameter estimates and slightly different standard error estimates. By default (i.e., without `se = "boot"`), `lavaan` calculates the standard error, z-value, and p-value for the conditional indirect effects using the delta method (i.e., the method described in Table 2), but this method is well known to be overly conservative. By adding the `se = "boot"` option, percentile bootstrap confidence intervals are calculated and printed.

The `ci.lower` and `ci.upper` columns provide the lower and upper confidence limits of a percentile bootstrap confidence interval for each of the conditional indirect effects and the index of moderated mediation. The confidence intervals do not exactly match the output from MEMORE; however, this is due to a difference in random seed used for each software, rather then differences in methods.

There are a number of options in this function which can be adjusted to align with options in `MEMORE`. For example, the level of the confidence interval can be changed using the `level` argument. The default is 0.95, but if a researcher wanted a 90% CI instead, they could add to the above command `level = 0.90` and this would adjust the level of the CI. Similarly, the number of bootstraps can be controlled with the `bootstrap` argument.

MEMORE provides additional output to what has been displayed in `lavaan`. For example, R^2 for each model is printed in MEMORE. This can also be print in `lavaan` by adding `rsquare = TRUE` to the command above.

Observant readers may have noticed that the total effect model was not estimated in the above `lavaan` model. This is because `lavaan` (and most other SEM programs) do not allow for simultaneous estimation of two equations for the same outcome. The total effect can be estimated in a separate model or the components of the total effect model can be calculated from the original model.

A separate model is estimated below (Equivalent to Equation 33 in the main manuscript). The conditional effect of X on Y at the two observed values of W (0 = delayed treatment, 1 = immediate treatment) are new parameters named `total0` $= \theta_{X\to Y}(W = 0)$ and `total1` $= \theta_{X\to Y}(W = 1)$.

```
BellTotalModel <- '
Ydiff ~ c0*1+c1*cond

total0 := c0
total1 := c0+c1
'

BellTotalFit <- sem(BellTotalModel, data = BellData)
summary(BellTotalFit, rsquare = TRUE, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 8 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                         3

  Number of observations                            99

Model Test User Model:

  Test statistic                                 0.000
  Degrees of freedom                                 0

Parameter Estimates:
```

```
  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Ydiff ~
    cond      (c1)   -0.036    1.007   -0.035    0.972   -2.009    1.938

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Ydiff     (c0)    0.017    0.708    0.024    0.981   -1.371    1.405

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Ydiff            25.091    3.566    7.036    0.000   18.101   32.081

R-Square:
                   Estimate
    Ydiff             0.000

Defined Parameters:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    total0            0.017    0.708    0.024    0.981   -1.371    1.405
    total1           -0.018    0.716   -0.026    0.979   -1.421    1.384
```

Compared to the results from MEMORE (using OLS), the parameter estimates are identical but the standard errors of the coefficients are slightly different due to the asympototic variance estimate used in maximum likelihood as compared to the finite sample variance estimate in OLS. The patterns of significance are no different.

Also not in the current output is the conditional direct effect. These can also be estimated as part of the lavaan model through some creative use of the tool. The conditional direct effects are conditional on the value of the moderator, but also the value of the mean-centered average of the mediators that is expected at that value of the moderator. This can be achieved by adding a model for the mean-centered average of the moderators and using these parameters to calculate the conditional direct effects. Indeed, this is also what MEMORE does in the background to calculate these effects.

```
BellDirectModel <- '
Mavg ~ h0*1+h1*cond

```

```
Mavg0 := h0
Mavg1 := h0+h1

direct0 := cp0+d1*Mavg0
direct1 := cp0+cp1+d1*Mavg1
'
```

This additional text can be combined with the original model syntax to estimate the conditional direct effects

```
BellDirectFit <- sem(c(BellModel, BellDirectModel), data = BellData)
summary(BellDirectFit, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 1 iteration

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        11

  Number of observations                            99

Model Test User Model:

  Test statistic                                 8.183
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.004

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Mdiff ~
    cond      (a1)   10.635    4.626    2.299    0.021    1.569   19.701
  Ydiff ~
    cond     (cp1)   -0.478    1.035   -0.462    0.644   -2.508    1.551
    Mdiff     (b1)    0.047    0.021    2.188    0.029    0.005    0.089
    Mavg      (d1)   -0.008    0.034   -0.244    0.807   -0.074    0.058
  Mavg ~
```

```
    cond      (h1)    6.686    2.928    2.283    0.022    0.947   12.424

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff     (a0)    5.728    3.254    1.760    0.078   -0.650   12.106
   .Ydiff    (cp0)   -0.278    0.712   -0.391    0.696   -1.674    1.118
   .Mavg      (h0)   -3.309    2.060   -1.606    0.108   -7.346    0.728

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff           529.516   75.262    7.036    0.000  382.005  677.027
   .Ydiff            23.990    3.410    7.036    0.000   17.307   30.673
   .Mavg            212.171   30.157    7.036    0.000  153.065  271.277

Defined Parameters:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
    indirect0         0.268    0.195    1.371    0.170   -0.115    0.651
    indirect1         0.766    0.382    2.003    0.045    0.016    1.515
    IMM               0.498    0.314    1.585    0.113   -0.118    1.113
    Mavg0            -3.309    2.060   -1.606    0.108   -7.346    0.728
    Mavg1             3.377    2.081    1.623    0.105   -0.702    7.455
    direct0          -0.251    0.704   -0.357    0.721   -1.630    1.128
    direct1          -0.784    0.783   -1.002    0.316   -2.318    0.750
```

Notice that the `Mavg0` and `Mavg1` parameter estimates match the values in the Mavg column for the direct effects in the MEMORE output. Similarly the direct effect estimates match exactly, but again because of slight differences in standard error estimation, there are slight differences in SE, and p-values.

As previously demonstrated, the total effect can be estimated as a separate model, and the conditional total effects can be estimated in that model. However, this can also be achieved by adding together the conditional direct and indirect effects, demonstrated below

```
BellTotalCalcModel <- '
total0 := direct0+indirect0
total1 := direct1+indirect1
'

BellTotalCalcFit <- sem(c(BellModel, BellDirectModel, BellTotalCalcModel),
data = BellData, se  =  "bootstrap", bootstrap  =  5000)
summary(BellTotalCalcFit, ci = TRUE)
```

```
lavaan 0.6.17 ended normally after 1 iteration
```

```
  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        11

  Number of observations                            99

Model Test User Model:

  Test statistic                                 8.183
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.004

Parameter Estimates:

  Standard errors                            Bootstrap
  Number of requested bootstrap draws             5000
  Number of successful bootstrap draws            5000

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
  Mdiff ~
    cond      (a1)   10.635    4.668    2.278    0.023    1.379   19.629
  Ydiff ~
    cond     (cp1)   -0.478    0.992   -0.482    0.630   -2.391    1.421
    Mdiff     (b1)    0.047    0.022    2.151    0.031    0.003    0.088
    Mavg      (d1)   -0.008    0.038   -0.216    0.829   -0.087    0.064
  Mavg ~
    cond      (h1)    6.686    2.941    2.273    0.023    0.977   12.508

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff     (a0)    5.728    2.390    2.396    0.017    1.104   10.529
   .Ydiff    (cp0)   -0.278    0.682   -0.408    0.683   -1.634    1.091
   .Mavg      (h0)   -3.309    1.781   -1.857    0.063   -6.723    0.285

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
   .Mdiff           529.516  110.722    4.782    0.000  328.611  755.065
   .Ydiff            23.990    4.085    5.873    0.000   15.336   31.498
   .Mavg            212.171   32.438    6.541    0.000  146.155  272.517

Defined Parameters:
```

```
             Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper
indirect0       0.268    0.175    1.535    0.125   -0.000    0.668
indirect1       0.766    0.417    1.838    0.066    0.042    1.689
IMM             0.498    0.336    1.480    0.139   -0.018    1.291
Mavg0          -3.309    1.782   -1.857    0.063   -6.723    0.285
Mavg1           3.377    2.336    1.445    0.148   -1.122    8.072
direct0        -0.251    0.655   -0.383    0.702   -1.538    1.082
direct1        -0.784    0.696   -1.127    0.260   -2.131    0.600
total0          0.017    0.610    0.028    0.978   -1.132    1.272
total1         -0.018    0.807   -0.023    0.982   -1.583    1.581
```

Notice that the results are identical which ever way the conditional total effects are calculated. The bootstrap confidence intervals for indirect effects and the index of moderated mediation are the same as previously calculated (the added code does not change these parameter estimates).

This provides all of the relevant output for conducting a first-stage conditional process analysis with a moderated direct effect in R using `lavaan`. It is worthwhile to note that the multiple steps taken in this tutorial are meant to be instructive, but the final model (`BellTotalCalcFit`) provides all of the information needed and users need not calculate each model separately and incrementally. This is done merely as a pedagogical activity.

Appendix C

Example 1: Lasselin et al. (2016)

Output from the MEMORE Macro. This output was generated by the SPSS version of the MEMORE macro when the command

```
MEMORE y = PostPain PrePain /m = PostPCS PrePCS /w = inflame /model = 16
        /seed = 1169404.
```

was applied to the data from the Lasselin et al. (2016) study.

```
Run MATRIX procedure:


*********************** MEMORE Procedure for SPSS Version 3.0 ************************
                          Written by Amanda Montoya
                   Documentation available at akmontoya.com
**************************************************************************************
Model:
  16
Variables:
Y = PostPain PrePain
W = inflame
M = PostPCS  PrePCS


Computed Variables:
Ydiff =          PostPain  -       PrePain
Mdiff =          PostPCS   -       PrePCS
Mavg =  (        PostPCS   +       PrePCS   )        /2                          Centered
Int1  = inflame  *        (        PostPCS   -       PrePCS   )
Sample Size:
  78
Seed:
  1169404
**************************************************************************************
```

```
Outcome: Ydiff =  PostPain  -       PrePain

Model Summary

          R       R-sq        MSE          F        df1        df2          p

      .3309      .1095     1.3808     9.3450     1.0000    76.0000      .0031

Model

             Effect         SE          t          p       LLCI       ULCI

constant     -.2310      .1331    -1.7363      .0866     -.4960      .0339

W             .4016      .1314     3.0570      .0031      .1400      .6633

Degrees of freedom for all regression coefficient estimates:

  76

Conditional Effect of Focal Predictor on Outcome at values of Moderator(s)

 Focal:   'X'      (X)

 Outcome: Ydiff    (Y)

 Mod:     inflame  (W)

    inflame     Effect         SE          t          p       LLCI       ULCI

    -1.0186     -.6401      .1888    -3.3910      .0011    -1.0161     -.2642

      .0006     -.2308      .1331    -1.7344      .0869     -.4958      .0342

     1.0199      .1786      .1888      .9461      .3471     -.1974      .5546

     Values for quantitative moderators are the mean and plus/minus one SD from the mean.

Degrees of freedom for all conditional effects:

  76

**************************************************************************************

Outcome: Mdiff =  PostPCS   -       PrePCS

Model

             Effect         SE          t          p       LLCI       ULCI

constant     5.0658      .9654     5.2472      .0000     3.1434     6.9882

Degrees of freedom for all regression coefficient estimates:

  77

**************************************************************************************

Outcome: Ydiff =  PostPain  -       PrePain

Model Summary

          R       R-sq        MSE          F        df1        df2          p

      .3412      .1164     1.4071     3.2508     3.0000    74.0000      .0265
```

Model

	coeff	SE	t	p	LLCI	ULCI
constant	-.2429	.1623	-1.4961	.1389	-.5663	.0806
Mdiff	.0079	.0192	.4116	.6818	-.0304	.0462
Int1	.0334	.0161	2.0791	.0411	.0014	.0655
Mavg	-.0211	.0120	-1.7597	.0826	-.0451	.0028

Degrees of freedom for all regression coefficient estimates:

74

Conditional Effect of Focal Predictor on Outcome at values of Moderator(s)

Focal: Mdiff (M)

Outcome: Ydiff (Y)

Mod: inflame (W)

inflame	Effect	SE	t	p	LLCI	ULCI
-1.0186	-.0261	.0176	-1.4805	.1430	-.0613	.0090
.0006	.0079	.0192	.4126	.6811	-.0304	.0463
1.0199	.0420	.0311	1.3518	.1806	-.0199	.1039

Values for quantitative moderators are the mean and plus/minus one SD from the mean.

Degrees of freedom for all conditional effects:

74

******************* CONDITIONAL TOTAL, DIRECT, AND INDIRECT EFFECTS *******************

Conditional Total Effect of X on Y at values of the Moderator(s)

inflame	Effect	SE	t	df	p	LLCI	ULCI
-1.0186	-.6401	.1888	-3.3910	76.0000	.0011	-1.0161	-.2642
.0006	-.2308	.1331	-1.7344	76.0000	.0869	-.4958	.0342
1.0199	.1786	.1888	.9461	76.0000	.3471	-.1974	.5546

Values for quantitative moderators are the mean and plus/minus one SD from the mean.

Direct effect of X on Y

Effect	SE	t	df	p	LLCI	ULCI
-.2429	.1623	-1.4961	74.0000	.1389	-.5663	.0806

Conditional Indirect Effect of X on Y through Mediator at values of the Moderator

Ind: Ind1

Med: Mdiff (M)

inflame	Effect	BootSE	BootLLCI	BootULCI

```
    -1.0186     -.1324      .0637     -.2637     -.0134
      .0006      .0402      .0693     -.0929      .1820
     1.0199      .2128      .1133      .0008      .4536
     Values for quantitative moderators are the mean and plus/minus one SD from the mean.
Indirect Key
Ind1  'X'      ->       Mdiff    ->       Ydiff
******************************** INDICES OF MODERATION ********************************
Test of Moderation of the Total Effect
      Effect         SE          t         df          p       LLCI       ULCI
W      .4016      .1314     3.0570    76.0000      .0031      .1400      .6633
Index of Moderated Mediation for each Indirect Effect.
          Effect     BootSE   BootLLCI   BootULCI
Ind1       .1693      .0593      .0558      .2911
**************************** ANALYSIS NOTES AND WARNINGS *****************************
NOTE: Some cases were deleted due to missing data. The number of cases was:
  2
Bootstrap confidence interval method used: Percentile bootstrap.
Number of bootstrap samples for bootstrap confidence intervals:
  5000
The following variables were mean centered prior to analysis:
 (        PostPCS   +       PrePCS   )        /2
Level of confidence for all confidence intervals in output:
      95.00
------ END MATRIX -----
```

Appendix D

Example 2: Bell et al., (2017)

Output from the MEMORE Macro. This output was generated by the SPSS version of the MEMORE macro when the command

```
MEMORE y = PostPEP PrePEP /m = PostPos PrePos /w = cond
/model = 12 /plot = 1
/seed = 104874.
```

was applied to the data from the Bell et al. (2017) study.

```
Run MATRIX procedure:


*********************** MEMORE Procedure for SPSS Version 3.0 ************************
                           Written by Amanda Montoya
                    Documentation available at akmontoya.com
**************************************************************************************
Model:
  12
Variables:
Y = PostPEP  PrePEP
W = cond
M = PostPos  PrePos
Computed Variables:
Ydiff =          PostPEP   -       PrePEP
Mdiff =          PostPos   -       PrePos
Mavg =  (        PostPos   +       PrePos   )        /2                          Centered
Sample Size:
  99
Seed:
  104874
**************************************************************************************
Outcome: Ydiff =  PostPEP   -       PrePEP
```

```
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .0035      .0000    25.6087      .0012     1.0000    97.0000      .9722
Model
             Effect          SE           t           p       LLCI       ULCI
constant      .0170       .7157       .0238       .9811    -1.4034     1.4374
W            -.0355      1.0173      -.0349       .9722    -2.0545     1.9835
Degrees of freedom for all regression coefficient estimates:
  97
Conditional Effect of Focal Predictor on Outcome at values of Moderator(s)
 Focal:   'X'      (X)
 Outcome: Ydiff    (Y)
 Mod:     cond     (W)
     cond       Effect          SE           t           p       LLCI       ULCI
      .0000      .0170       .7157       .0238       .9811    -1.4034     1.4374
     1.0000     -.0185       .7229      -.0256       .9796    -1.4533     1.4163
     Values for dichotomous moderators are the two values of the moderator.
Degrees of freedom for all conditional effects:
  97
**************************************************************************************
Outcome: Mdiff =  PostPos   -       PrePos
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2251      .0507   540.4336     5.1795     1.0000    97.0000      .0251
Model
             Effect          SE           t           p       LLCI       ULCI
constant     5.7280      3.2877      1.7423       .0846     -.7971    12.2531
W           10.6353      4.6731      2.2758       .0251     1.3604    19.9101
Degrees of freedom for all regression coefficient estimates:
  97
Conditional Effect of Focal Predictor on Outcome at values of Moderator(s)
 Focal:   'X'      (X)
 Outcome: Mdiff    (M)
```

```
 Mod:     cond     (W)
     cond       Effect          SE           t           p        LLCI        ULCI
      .0000     5.7280      3.2877      1.7423       .0846      -.7971     12.2531
     1.0000    16.3633      3.3210      4.9272       .0000      9.7719     22.9546
     Values for dichotomous moderators are the two values of the moderator.
Degrees of freedom for all conditional effects:
  97
**************************************************************************************
Outcome: Ydiff =  PostPEP   -       PrePEP
Model Summary
          R       R-sq         MSE           F         df1         df2           p
      .2095      .0439     25.0004      1.4538      3.0000     95.0000       .2321
Model
             coeff          SE           t           p        LLCI        ULCI
constant     -.2783       .7348      -.3788       .7057     -1.7371      1.1804
W            -.4780      1.0459      -.4571       .6487     -2.5543      1.5983
Mdiff         .0468       .0228      2.0561       .0425       .0016       .0920
Mavg         -.0083       .0360      -.2295       .8189      -.0796       .0631
Degrees of freedom for all regression coefficient estimates:
  95
--------------------------------------------------------------------------------------
Conditional Effect of Focal Predictor on Outcome at values of Moderator(s)
 Focal:   'X'      (X)
 Outcome: Ydiff    (M)
 Mod:     cond     (W)
     cond         Mavg     Effect          SE           t           p        LLCI        ULCI
      .0000    -3.3090     -.2510       .7190      -.3491       .7278     -1.6785      1.1764
     1.0000     3.3765     -.7842       .8055      -.9735       .3328     -2.3834       .8150
     Values for dichotomous moderators are the two values of the moderator.
     Values for mediator averages are the conditional values based on the values of the moderator.
--------------------------------------------------------------------------------------
Degrees of freedom for all conditional effects:
  95
```

```
****************** CONDITIONAL TOTAL, DIRECT, AND INDIRECT EFFECTS ******************

Conditional Total Effect of X on Y at values of the Moderator(s)

     cond     Effect        SE         t        df         p      LLCI      ULCI

    .0000      .0170     .7157     .0238   97.0000     .9811   -1.4034    1.4374

   1.0000     -.0185     .7229    -.0256   97.0000     .9796   -1.4533    1.4163

     Values for dichotomous moderators are the two values of the moderator.

Conditional Direct Effect of X on Y at values of the Moderator(s)

     cond       Mavg    Effect        SE         t        df         p      LLCI      ULCI

    .0000    -3.3090    -.2510     .7190    -.3491   95.0000     .7278   -1.6785    1.1764

   1.0000     3.3765    -.7842     .8055    -.9735   95.0000     .3328   -2.3834     .8150

     Values for dichotomous moderators are the two values of the moderator.

     Values for mediator averages (Mavg) are the conditional values based on the values

     of the moderator.

Conditional Indirect Effect of X on Y through Mediator at values of the Moderator

  Ind:      Ind1

  Med:      Mdiff     (M)

     cond     Effect    BootSE  BootLLCI  BootULCI

    .0000      .2680     .1750    -.0062     .6700

   1.0000      .7657     .4257     .0401    1.6979

     Values for dichotomous moderators are the two values of the moderator.

Indirect Key

Ind1  'X'     ->      Mdiff    ->      Ydiff

******************************* INDICES OF MODERATION *******************************

Test of Moderation of the Total Effect

      Effect        SE         t        df         p      LLCI      ULCI

W     -.0355    1.0173    -.0349   97.0000     .9722   -2.0545    1.9835

Test of Moderation of the Direct Effect

      Effect        SE         t        df         p      LLCI      ULCI

W     -.4780    1.0459    -.4571   95.0000     .6487   -2.5543    1.5983

Index of Moderated Mediation for each Indirect Effect.

          Effect    BootSE  BootLLCI  BootULCI

Ind1       .4977     .3445    -.0243    1.3096
```

```
************************************ PLOTS *******************************************
Data for visualizing conditional effect of X on Y at values of W
Data for visualizing conditional effect of X on M at values of W
Paste text below into a SPSS syntax window and execute to produce plot.
DATA LIST FREE/cond YdiffHAT MdiffHAT.
BEGIN DATA.

      .0000      .0170     5.7280
     1.0000     -.0185    16.3633
END DATA.
GRAPH/SCATTERPLOT = cond WITH YdiffHAT.
GRAPH/SCATTERPLOT = cond WITH MdiffHAT.
--------------------------------------------------------------------------------------
Data for visualizing conditional direct effect of X on Y at values of W.
Paste text below into a SPSS syntax window and execute to produce plot.
Note: All mediator averages are conditioned on their predicted value based on W.
DATA LIST FREE/cond YdiffHAT.
BEGIN DATA.
      .0000     -.2510
     1.0000     -.7842
END DATA.
GRAPH/SCATTERPLOT = cond WITH YdiffHAT.
*************************** ANALYSIS NOTES AND WARNINGS ****************************
Bootstrap confidence interval method used: Percentile bootstrap.
Number of bootstrap samples for bootstrap confidence intervals:
  5000
The following variables were mean centered prior to analysis:
 (        PostPos   +       PrePos   )        /2
Level of confidence for all confidence intervals in output:
      95.00
------ END MATRIX -----
```