Published in final edited form as:

Ann Nutr Metab. 2014 ; 65(2-3): 198–204. doi:10.1159/000362505.

Mediation from Multilevel to Structural Equation Modeling

David P. MacKinnon and Matthew J. Valente

Department of Psychology, Arizona State University, Tempe, Ariz., USA

Abstract

Background/Aims—The purpose of this article is to outline multilevel structural equation modeling (MSEM) for mediation analysis of longitudinal data. The introduction of mediating variables can improve experimental and nonexperimental studies of child growth in several ways as discussed throughout this article. Single-mediator individual-level and multilevel mediation models illustrate several current issues in the estimation of mediation with longitudinal data. The strengths of incorporating structural equation modeling (SEM) with multilevel mediation modeling are described.

Summary and Key Messages—Longitudinal mediation models are pervasive in many areas of research including child growth. Longitudinal mediation models are ideally modeled as repeated measurements clustered within individuals. Further, the combination of MSEM and SEM provides an ideal approach for several reasons, including the ability to assess effects at different levels of analysis, incorporation of measurement error and possible random effects that vary across individuals.

Keywords

Mediation; Multilevel mediation; Multilevel structural equation modeling

Introduction

Statistical analyses are often conducted to understand relations between two variables, such as the relation between age and size. As a research area matures, research questions are elaborated to address how third variables affect bivariate relationships, such as how nutrition affects the relationship between age and size. The addition of a third variable into theoretical models can result in a variety of effects known as third-variable effects. Third variables can be described as moderators, confounders, covariates and mediators. The focus of this article is on mediator variables. A mediator is a variable (M) that affects a dependent variable and is itself affected by an independent variable [1] and should be contrasted with a confounding variable which is related to two variables but is not in a causal mediation sequence between the two variables. Mediators are unique because these variables transmit part or all of the

Prof. David P. MacKinnon, PhD, Department of Psychology, Arizona State University, PO Box 871104, 950 South McAllister Street, Room 237, Tempe, AZ 85287-1104 (USA), David.MacKinnon@asu.edu.

This paper was presented at the 2013 Child Growth Trajectory Workshop in Munich, Germany.

Disclosure Statement

The authors have no conflict of interest.

influence of an independent variable (X) to a dependent variable (Y) [2, 3]. Statistical mediation analysis is used to assess evidence of a mediation process that describes how, or through what mechanism, an independent variable is related to a dependent variable. Investigating how an independent variable is related to a dependent variable is important for isolating and understanding sequences of causal events and is an important question in many disciplines [1, 4, 5], including child growth.

Mediating processes are hypothesized for many aspect of child growth. For example, recently, a randomized clinical trial was conducted to determine the effects of protein content of infant formulas on children's body mass index (BMI) later in life [6]. In this study, children were randomly assigned to conditions for which they were either fed high- or low-protein infant formula for the first 6 months of their lives. At 6 months of age, insulin-like growth factor-1 axis serum levels were assessed and various other outcomes were measured at 2 and 6 years of age. Of main importance are insulin-like growth factor-1 axis serum level and BMI outcome measures. Protein intake of children during the first 6 months of their lives (X) resulted in higher insulin-like growth factor-1 axis serum levels at 6 months of age (M) which in turn resulted in higher BMI levels at 6 years of age (Y). The area of child growth has many examples of hypothesized mediating processes [7–10]. For example, girls' weight status (X) is related to math performance (Y) through its relationship with interpersonal skills (M) [7]. Iron deficiency anemia (X) in children is related to levels of physical activity (Y) through its relationship with cognitive development (M) [9]. Intervention programs (X) promote healthy behavior (Y) by changing norms (M) that are associated with those behaviors [10].

There are two overlapping motivations for the study of mediating variables. First, statistical mediation can be planned before an experiment and incorporated into an experimental design. In this case, a mediator is targeted specifically by an experimental manipulation because it is hypothesized to be causally related to an outcome variable [4]. If a mediating variable is causally related to a dependent variable and an experiment is conducted that changes the mediator, the experimental manipulation will change the dependent variable. In this mediation by design example, the mediating variables are selected before the intervention. An example of mediation by design would be an experimental manipulation to change nutrition that is hypothesized to then affect child growth. A second use of mediation analysis is to explain an observed relationship between two variables. Mediation analysis is a way to explain what mechanism was responsible for an observed relationship between an independent variable and a dependent variable. Mediation for explanation generally investigates mediation after an observed relation is found. Statistical mediation by design relies on a prior hypothesis regarding the role of a mediating variable and statistical mediation for explanation relies on post hoc explanations of how two variables were related within a study.

The Single-Mediator Model

The single-mediator model is described first and models that include longitudinal measurement are described later. The following three regression equations are used to

represent the relations between a single independent variable (X), a single mediating variable (M) and a single dependent variable (Y):

$$\hat{Y} = \hat{\alpha}_1 + \hat{c}X + e_1 \quad (1)$$

$$\hat{Y} = \hat{\alpha}_2 + \hat{c}'X + \hat{b}M + e_2 \quad (2)$$

$$\hat{M} = \hat{\alpha}_3 + \hat{a}X + e_3 \quad (3)$$

Equation 1 represents the relation of an independent variable (X) to a dependent variable (Y) through the coefficient $\hat{c}$. Equation 2 represents the relation between an independent variable (X) and a dependent variable (Y) adjusting for a mediating variable (M) through the coefficient $\hat{c}'$ and it represents the relation of a mediating variable (M) to a dependent variable (Y) adjusting for an independent variable (X) through the coefficient $\hat{b}$. Equation 3 represents the relation between an independent variable (X) and a mediating variable (M) through the coefficient $\hat{a}$. Intercepts in equations 1–3 are represented by $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$, respectively, and e_1 , e_2 and e_3 represent the unexplained variability in Y after accounting for the predictors in equations 1–3, respectively. Though not shown in equation 2, an interaction of X and M on Y can be added to equation 2 to investigate whether the relation of M to Y differs across levels of X . These equations can be estimated using multiple regression or structural equation modeling (SEM) software. Often only equations 2 and 3 are estimated to investigate mediation [11].

The Mediated Effect

There are different approaches to estimating mediated effects from regression models. The product of the $\hat{a}$ and $\hat{b}$ coefficients ($\hat{a}\hat{b}$) is an estimator of the mediated effect under certain assumptions, including lack of confounding of the X to M and M to Y relations. Because X affects Y indirectly through M , the mediated effect is also known as the *indirect effect*. The effect of X on Y after adjustment for M ($\hat{c}'$) is an estimator of the *direct effect*. The mediated effect is also equal to the difference between the $\hat{c}$ and $\hat{c}'$ coefficients ($\hat{c} - \hat{c}'$). The rationale behind the $\hat{a}\hat{b}$ mediation measure is that mediation depends on the extent to which the independent variable affects the mediator (coefficient $\hat{a}$) and the extent to which the mediator affects the dependent variable (coefficient $\hat{b}$). Similarly, the change in the $\hat{c}$ coefficient when adjusted for the mediator, $\hat{c}'$ reflects how much of the relation between the independent variable and the dependent variable is explained by the mediator. As a result, the *total effect* $\hat{c}$ (fig. 1, top panel) can be decomposed into a direct effect $\hat{c}'$ and an indirect effect $\hat{a}\hat{b}$ (fig. 1, bottom panel). For the multiple regression equations described above, $\hat{a}\hat{b} = \hat{c} - \hat{c}'$, but these quantities are not equal for other models such as logistic regression [12].

Recent developments in causal mediation analysis have generated estimators of mediation effects accurate for linear and nonlinear models [13]. For accurate identification of mediation effects, several assumptions are needed regarding the causal relations among variables for identification of effects [14]: (1) no confounders of the X to M relation, (2) no confounders of the M to Y relation, (3) no confounders of the X to Y relation and (4) no

effects of X on confounders that then affect the relation of X to Y . Randomization of X is assumed so that assumptions 1 and 3 are satisfied which allows for estimators $\hat{a}$ and $\hat{c}$ to be treated as causal effects. Randomization of X does not satisfy assumptions 2 and 4 which are required for interpretation of $\hat{b}$ and $\hat{c}'$ as causal effects. Violation of these assumptions can be addressed with statistical methods that adjust for confounders and also methods that assess the sensitivity of results to violation of these assumptions [15]. However, randomization of X does not guarantee that the mediator has been randomized to persons and is a general weakness of the mediation model as now well described in the research literature [4, 13, 14]. Furthermore, for any longitudinal mediation analysis, theory for when effects occur for X to M and M to Y must match the times when data are collected to obtain accurate estimates of effects [4]. Longitudinal issues may be more complicated for multilevel analysis because if M and Y are measured concurrently, there is no lag, yet temporal precedence is necessary for causal inference in mediation models at least in terms of the causal ordering of M and Y , when X is randomized. Even if M is lagged relative to Y when estimating the effects, the between-cluster effects still represent the averages across concurrent measurements.

Multilevel Mediation Models

Equations 2 and 3 above can be rewritten to include repeated measures among individuals for M and Y as shown below in equations 4 and 5 where X represents randomization to experimental conditions and M_1/M_2 and Y_1/Y_2 represent repeated measures of M and Y , respectively. Equations 4 and 5 can be interpreted similarly to equations 2 and 3 but with additional coefficients representing stability of Y measured across times 1 and 2 ($\hat{S}_y$) and stability of M measured across times 1 and 2 ($\hat{S}_m$) added into the equations (fig. 2).

$$\hat{Y}_2 = \hat{\alpha}_4 + \hat{c}'X + \hat{S}_y Y_1 + \hat{b}M_2 + e_4 \quad (4)$$

$$\hat{M}_2 = \hat{\alpha}_5 + \hat{a}X + \hat{S}_m M_1 + e_5 \quad (5)$$

Equations 6, 7, 9, and 10 specify multilevel models for within-individual measures (level 1) and between-individual measures (level 2) corresponding to equations 2 and 3 for the single-level analysis and equations 4 and 5 for the repeated-measure example. The independent variable is at the between-individual level, and the mediator and the dependent variables are at the within-individual level. At level 1, a model is specified for repeated observations within each individual that could be for two observations per person as in equations 4 and 5 or for more observations per person. At level 2, another linear model is specified, but the dependent variable is the intercept (and possibly slopes, although only random intercepts are shown in these equations) in the level-1 model. As a result, there are within- and between-individual equations. In these equations, i subscripts refer to time, j subscripts refer to individuals and Y and M refer to the dependent and mediating variable, respectively.

Multilevel Equations for Y Predicted by X and M

$$\text{Within-individual level 1: } \hat{Y}_{ijY} = \hat{B}_{0jY} + \hat{b}M_{ijY} + e_{ijY} \quad (6)$$

$$\text{Between-individual level 2: } \hat{B}_{0jY} = \hat{\gamma}_{00Y} + \hat{c}'X_{jY} + u_{0jY} \quad (7)$$

Equations 6 and 7 include two predictors, one at the within-individual level, M , and the other at the between-individual level, X . In equation 6, the within-individual level score on the dependent variable is equal to an intercept $\hat{B}_{0jY}$ plus random error, e_{ijY} associated with the i th score for the j th individual for the prediction of Y . The within-level random error, e_{ijY} , is assumed to have a normal distribution with $\text{Var}(e_{ijY}) = \Phi_Y^2$. In equation 7, the dependent variable is the between-level intercept, $\hat{B}_{0jY}$, which is equal to an overall mean, $\hat{\gamma}_{00Y}$, plus the slope, $\hat{c}'$, times independent variable X , plus the residual representing the difference of the cluster mean and the grand mean, u_{0jY} . The b estimate is at the within-level, because the mediator is assumed to work through within-person processes. The c' estimate, on the other hand, is in the between-level equation because the individuals are assigned to conditions for this example. There is some controversy regarding the estimation of mediation by using coefficients at different levels, e.g. the X to M relation at the between level and the M to Y relation at the within level related to the theoretical mediation model [16, 17].

There are several additional models that include the X and M predictors that may be appropriate given the substantive context of the research. One such model investigates whether the slopes relating M to Y differ across the individuals. This model would include two level-2 regression equations; one for the random slope and one for the random intercept, so in principle b in equation 6 could be random. Another model would include the between-level mean of M as an additional predictor to investigate both between- and within-level relations between M and Y . The deviations in the level-2 equation, u_{0jY} , are assumed to have a normal distribution with variation between individual means, $\text{Var}(u_{0jY}) = \vartheta_{00Y}$. The c' coefficient is estimated at the between-level because assignment to a condition is assumed to be at the between-level for this example. The estimation of error terms at both levels of the model (e_{ijY} at the within-individual level and u_{0jY} at the between-individual level) allows for a non-zero intraclass correlation (ICC) to be incorporated in the analysis. As shown in equation 8, random-effect estimates of the variance between individuals, ϑ_{00Y} and the variance of the residuals at the within-individual level, Φ_Y^2 provide an estimator of the residual ICC.

$$\text{ICC} = \vartheta_{00Y} / (\vartheta_{00Y} + \Phi_Y^2) \quad (8)$$

Using this equation, the ICC conditional on other effects in the model can be easily calculated. The residual ICC can be used to assess the amount of dependency after adjusting for different predictors.

Multilevel Equations for M Predicted by X

$$\text{Within-individual level 1: } \hat{M}_{ij} = \hat{B}_{0jM} + e_{ijM} \quad (9)$$

$$\text{Between-individual level 2: } \hat{B}_{0jM} = \hat{\gamma}_{00M} + \hat{a}X_{jM} + u_{0jM} \quad (10)$$

Equations 9 and 10 are analogous to equation 3, but X predicts the dependent variable M rather than Y in a multilevel framework. The $\hat{a}$ estimate at the between-level because the assignment to conditions is at the between-level for this example.

Because of the complex structure of the multilevel model with error terms at multiple levels, the parameters of the model are estimated with iterative methods such as restricted maximum likelihood techniques, rather than the ordinary least squares (OLS) methods typically employed to estimate the parameters of single-level models [18–20]. The standard error estimates for the multilevel model are consequently more accurate than those for a single-level individual-as-unit-of-analysis model of multilevel data because they incorporate the dependence of subjects measured within groups (i.e. a non-zero ICC).

The $\hat{a}\hat{b}$ and $\hat{c}-\hat{d}$ estimators of the mediated effect, algebraically equivalent in single-level models, are not exactly equivalent in the multilevel models [21]. The non-equivalence between $\hat{a}\hat{b}$ and $\hat{c}-\hat{d}$, however, is unlikely to be problematic because the discrepancy between the two estimates is typically small [21]. The standard error of the mediated effect estimates and standard errors of $\hat{a}$ and $\hat{b}$ may come from equations at different levels of analysis and may require the covariance between $\hat{a}$ and $\hat{b}$. As described below, there are situations where $\hat{a}$ and $\hat{b}$ are random effects and a different formula must be used to estimate the mediated effect [22].

Centering, which usually consists of removing the group mean from predictors, is important when analyzing both between- and within-level effects of the same predictor. In general, it is important to center predictor variables prior to estimation of multilevel models because the value and meaning of the intercept depends on the coding of the X variable and the intercept is a dependent variable in some equations. The intercept is the value of Y when all X variables are zero for any regression equation. If an X variable is not centered, then the intercept will be the value of Y when X is zero, even when a zero value of X is meaningless. After centering the X variables by subtracting the mean of the variable, the intercept is the value of Y at the average value of X . In the case of including both between- and within-level predictors, it is also important to create a new variable by subtracting the between level mean for each observation for a person. Making this new variable will simplify interpretation and reduce any correlation between the between- and within-level predictors [23–25].

There are different multilevel models corresponding to the level at which measurements are obtained as described by Krull and MacKinnon [26]. For example, the example equations described above correspond to a 2-1-1 model corresponding to X measured at the between level and M and Y measured at the within level. The 2-1-1 model is common when there is a randomized intervention, X , at the between level. If both X and M were measured at the second level, then the model would be a 2-2-1 model. If there are repeated measures for X , M and Y , then the model would be a 1-1-1 model. In this 1-1-1 model, it is possible to assess mediation effects at within-and between-individual levels.

There are several limitations of multilevel models as described by several researchers [27, 28]. Measurement error is not explicitly included in these models. Yet, many measures are obtained with error. Also, most discussions of multilevel mediation models focus on estimation equations separately [4], although there are multivariate multilevel models that could be applied [29]. There are also limited measures of model adequacy and fit in the multilevel mediation model. SEM are used to accommodate each of these limitations of the multilevel model. Several researchers have noted the need to combine the multilevel models with SEM to include the strengths of both approaches [30–32].

Multilevel Structural Equation Models

The mediation model is an SEM in which X causes M and M causes Y . SEM can be used to simultaneously estimate regression models [27, 28]. For example, SEM for mediation would simultaneously estimate equations 2 and 3 for the single-level model and equations 6, 7, 9, and 10 for the multilevel regression models. So the SEM allows for estimation of relations among variables at different levels of analysis. In this framework with two-level data, there are potentially two different models, one for the between-individual level data and a model for relations among times for an individual. Muthén and Muthén [33] represent these types of models as models for individual-level data (fig. 3, top panel) and models for the average among the individuals (i.e. cluster average), or, in the case of repeated measures, the person-specific average (fig. 3, bottom panel). In this case, there can be many different levels of analysis. Using the 1-1-1 example where repeated measures of X , M and Y are available, different models can be estimated at the within- and between-individual levels.

In summary, the multilevel structural equation model (MSEM) allows for measurement models for constructs thereby accommodating measurement error. MSEM provides a very general model that allows simultaneous estimation of all regression coefficients (e.g. mediation model coefficients) and some of the fit indices from SEM can be used to assess model fit in MSEM. Furthermore, there is software available to estimate a variety of MSEM including Mplus [33] and GLLAMM [32].

Example

To clarify, a few concepts consider a model with repeated measures of X , M and Y , e.g. repeated measures of available food (X), food intake (M) and height (Y). It is possible that the relation between X and M may differ across persons and the relation of M to Y , adjusted for X , may differ across persons. Also, the relation of X to Y adjusted for X may differ across persons. This would be a model with random slopes. In this situation, the covariance between $\hat{a}$ and $\hat{b}$ is needed to compute the mediated effect and the standard error of the mediated effect. In order to get this correlation or covariance, there are several options. Kenny et al. [22] noted the need to include this covariance and used a method similar to resampling to get the average correlation between $\hat{a}$ and $\hat{b}$. Bauer et al. [34] describe a method to compute the covariance using an estimation trick, and Mplus will estimate this correlation and covariance and use it in the estimation of standard errors of the mediated effect [4, 35].

One useful extension of MSEM is its use for growth curve modeling. SEM and multilevel models are equivalent ways to specify simple growth curves [36]. Multilevel modeling clearly includes the different levels and analyses at each level. Traditional SEM does not explicitly include the multiple levels but can specify latent variables to model measurement error. MSEM has the best of both methods but because of the options for levels of analysis and measurement error it can be tricky to program the exact model of interest [33]. Fortunately, there are several resources to help in this endeavor, including articles outlining the different modeling options [37] and new books to provide some guidance for these methods in mediation [4, 35].

Discussion

Mediation models are common in prevention and treatment studies in many fields [4] and are useful for the study of growth. Mediation analysis provides a way to investigate a chain of relations among variables, including testing the theory of how growth changes as a function of potential variables including intervention exposure. Although mediation analyses hold promise for uncovering causal mediation relations, there are limitations for how well the methods can identify effects in the presence of confounding variables. Designing randomized studies to change mediating variables improves interpretation of effects, but there can still be confounders of the M to Y relation because even though X is randomized, the value of the mediator is self-selected. Modern methods are available to investigate and adjust effects for possible confounding variables. Multilevel mediation models more accurately model mediation at different levels of analysis. The multilevel model can be specified as a special case of SEM. The combination of multilevel models for accurate investigation of relations at different levels of analysis with SEM for testing multiple equations simultaneously and incorporating measurement error makes MSEM a powerful technique that is only beginning to be applied. There are some complex issues for testing and evaluating these models, however, at least in part, because of the sheer number of potential models that can be tested. MSEM allows for incorporation of measurement error, relations among many variables including mediation and estimation of the entire model simultaneously. General software is available to estimate these models and there are resources to more clearly specify and test these models.

Acknowledgments

This research was supported in part by the National Institute on Drug Abuse (grant No. R01 DA09757).

References

1. Last, JM. A Dictionary of Epidemiology. 2. New York: Oxford University Press; 1988.
2. MacKinnon, DP. Mediating variable. In: Matsumoto, D., editor. The Cambridge Dictionary of Psychology. Cambridge: Cambridge University Press; 2009.
3. Sobel ME. Effect analysis and causation in linear structural equation models. *Psychometrika*. 1990; 55:495–515.
4. MacKinnon, DP. Introduction to Statistical Mediation Analysis. New York: Erlbaum; 2008.
5. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol*. 2007; 58:593–614. [PubMed: 16968208]

6. Rzehak P, Grote V, Lattka E, Weber M, Gruszfeld D, Socha P, Closa-Monasterolo R, Escribano J, Giovannini M, Verduci E, Goyens P, Martin F, Langhendries JP, Demmelmair H, Klopp N, Illig T, Koletzko B. European Childhood Obesity Trial Study Group. Associations of IGF-1 gene variants and milk protein intake with IGF-1 concentrations in infants at age 6 months – results from a randomized clinical trial. *Growth Horm IGF Res.* 2013; 23:149–158. [PubMed: 23800627]
7. Gable S, Krull JL, Chang Y. Boys' and girls' weight status and math performance from kindergarten entry through fifth grade: a mediated analysis. *Child Dev.* 2012; 83:1822–1839. [PubMed: 22694240]
8. Dziewaltowski DA, Estabrooks PA, Welk G, Hill J, Milliken G, Karteroliotis K, Johnston JA. Healthy youth places: a randomized controlled trial to determine the effectiveness of facilitating adult and youth leaders to promote physical activity and fruit and vegetable consumption in middle schools. *Health Educ Behav.* 2009; 36:583–600. [PubMed: 18469366]
9. Zemel B. Body composition during growth and development. In: Cameron, N., editor. *Human Growth and Development*. San Diego: Academic Press; 2002. p. 271-293.
10. Judd CM, Kenny DA. Process analysis: estimating mediation in treatment evaluations. *Eval Rev.* 1981; 5:602–619.
11. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods.* 2002; 7:83–104. [PubMed: 11928892]
12. MacKinnon DP, Lockwood CM, Brown CH, Wang W, Hoffman JM. The intermediate endpoint effect in logistic and probit regression. *Clin Trials.* 2007; 5:499–513. [PubMed: 17942466]
13. Pearl J. The causal mediation formula: a guide to the assessment of pathways and mechanisms. *Prev Sci.* 2012; 13:426–436. [PubMed: 22419385]
14. VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Stat Interface.* 2009; 2:457–468.
15. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods.* 2010; 15:309–334. [PubMed: 20954780]
16. Pituch KA, Stapleton LM. Distinguishing between cross- and cluster-level mediation processes in the cluster randomized trial. *Soc Methods Res.* 2012; 41:630–670.
17. VanderWeele TJ. Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Soc Methods Res.* 2010; 38:515–544.
18. Nezlek, JB. Multilevel modeling of diary-style data. In: Mehl, MR.; Conner, TS., editors. *Handbook of Research Methods for Studying Daily life*. New York: Guilford; 2012. p. 357-383.
19. Nezlek, JB. Multilevel modeling for psychologists. In: Cooper, H., editor. *APA Handbook of Research Methods in Psychology: Data Analysis and Research Publication*. Washington: American Psychological Association; 2012. p. 219-241.
20. Nezlek JB. Multilevel random coefficient analyses of event and interval contingent data in social and personality psychology research. *Pers Soc Psychol Bull.* 2001; 27:771–785.
21. Krull JL, MacKinnon DP. Multilevel mediation modeling in group based intervention studies. *Eval Rev.* 1999; 23:418–444. [PubMed: 10558394]
22. Kenny DA, Bolger N, Korchmaros JD. Lower-level mediation in multilevel models. *Psychol Methods.* 2003; 8:115–128. [PubMed: 12924810]
23. Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychol Methods.* 2007; 12:121–138. [PubMed: 17563168]
24. Kreft IGG, de Leeuw J, Aiken LS. The effect of different forms of centering in hierarchical linear models. *Multivariate Behav Res.* 1995; 30:1–21. [PubMed: 26828342]
25. Hofmann DA, Gavin MB. Centering decisions in hierarchical linear models: implications for research in organizations. *J Manag.* 1998; 24:623–641.
26. Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. *Multivariate Behav Res.* 2001; 36:249–277. [PubMed: 26822111]
27. Card, NA. Multilevel mediational analysis in the study of daily lives. In: Mehl, MR.; Conner, TS., editors. *Handbook of Research Methods for Studying Daily Life*. New York: Guilford; 2012. p. 479-493.

28. Rabe-Hesketh, S.; Skrondal, A.; Zheng, X. Multilevel structural equation modeling. In: Hoyle, RH., editor. *Handbook of Structural Equation Modeling*. New York: Guilford; 2012. p. 512-531.
29. Hox, JJ. *Multilevel Analysis: Techniques and Applications*. 2. Routledge, New York; 2010.
30. Heck, RH.; Thomas, SL. *An introduction to Multilevel Modeling Techniques*. 2. Routledge, New York; 2009.
31. Muthén BO. Multilevel covariance structure analysis. *Soc Methods Res*. 1994; 22:376–398.
32. Rabe-Hesketh, S.; Skrondal, A.; Pickles, A. *GLLAMM Manual*. Technical Report 160. UC Berkeley Division of Biostatistics Working Paper Series. 2004. <http://www.bepress.com/ucbbiostat/paper160/>
33. Muthén, LK.; Muthén, BO. *Mplus User's guide*. 7. Los Angeles: Muthén & Muthén; 2012.
34. Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychol Methods*. 2006; 11:142–163. [PubMed: 16784335]
35. Bolger, N.; Laurenceau, J-P. *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. New York: Guilford; 2013.
36. Chou, C-P.; Bentler, PM.; Pentz, MA. A two-stage approach to multilevel structural equation models: application to longitudinal data. In: Little, TD.; Schnabel, KU.; Baumert, J., editors. *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples*. Mahwah: Erlbaum; 2000. p. 33-49.
37. Preacher KJ, Zyphur MJ, Zhang Z. A general multilevel SEM framework for assessing multilevel mediation. *Psychol Methods*. 2010; 15:209–233. [PubMed: 20822249]

Fig. 1.

The top panel illustrates the total effect of independent variable X on dependent variable Y . The bottom panel illustrates the mediated effect of independent variable X on dependent variable Y through mediating variable M .

Longitudinal mediation

Fig. 2.

This figure illustrates the mediation of a randomized treatment variable X and repeated measures of mediating variable M and repeated measures of dependent variable Y .

Fig. 3.

The top panel illustrates a 1-1-1 model for which independent variable X , mediating variable M and dependent variable Y are measured across time (denoted with subscript i) and across individuals (denoted with subscript j). The bottom panel illustrates a 2-2-2 model for which X , M and Y are measured across individuals. Error terms across levels and equations are each unique.