

Multigroup Analysis and Moderation with SEM

Overview of How Group Differences Are Investigated in SEM

There are two general ways to investigate group differences with structural equation modeling (SEM). The first method follows the approach used in regression analysis in which a predictor is a binary (or set of dummy variables representing multiple categories) is a predictor of some other variable. For a simple path model when the outcome is a continuous, measured variable, the regression coefficient represents a test of group differences which is equal to a t -test or two-group Analysis of Variance (ANOVA) of the difference between two group means.¹

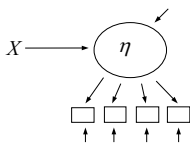


If the predictor is of course a set of dummy variables constructed to capture several nominal groups (e.g., religious categories), then the R-square test for variance accounted for in Y is equivalent to the multigroup ANOVA. The test of the regression coefficients for the set of paths represent comparisons to the referent group (coded 0 for all of the dummy variables). These equivalences mean that t -tests, ANOVA, and, of course, correlation and regression are special cases of SEM. Although I am not going to emphasize the point in this handout, when X and Y are both binary variables in the path model depicted above, we have a simple logistic (or probit) regression model, which tests the same hypothesis as a 2×2 chi-square. And extending this idea to the case with multiple dummy variables or ordinal or multicategory outcomes, one can use SEM to test a variety of hypotheses that are tested with traditional categorical tests (see Newsom, 2017, for more information).

The second general method of investigating group differences with SEM is to use multigroup models (Jöreskog, 1971; Sorböm, 1974). Multigroup models test separate models in two or more discrete groups. Equality constraints across groups are used to conduct nested tests using likelihood ratio comparisons between a model with certain parameters constrained to be equal and a model with those same parameters freely estimated (allowed to differ) across the groups. For example, one can investigate whether means, predictive paths, or loadings differ across two nationalities.

MIMIC Models

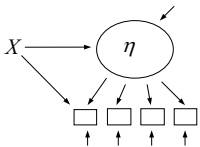
An extension of the simple regression path model depicted above has X as a binary predictor and a latent variable as the outcome.



This model tests the group differences in the latent variable and can be extended to multiple dummy variables of multiple predictors. The predictive path from X to η is a test of whether the two groups differ on the latent variable, where the use of a latent variable allows for the estimation of measurement error. This general approach to group differences is often referred to with by special term—the multiple indicator multiple cause (MIMIC) model (Jöreskog & Goldberger, 1975)—because it can include multiple predictors (“causes”) as well as multiple indicators of the latent variable. Use of this model can be valuable for examining research questions about differential item functioning, a concept from Item Response Theory (IRT; Lee, Little, & Preacher, 2012). If the group predictor variable (X) also has a path to one of the items, the path directly to the indicator variable provides information about group differences in the response that occur over and above the effects of the latent variable. For example, if

¹ The simple regression with a binary predictor is equivalent to a two-group analysis of variance or a correlation (sometimes given the name “point-biserial”). See the handout “ t -Tests, Chi-squares, Phi, Correlations: It’s all the same stuff” at my Univariate Statistics course page, <http://web.pdx.edu/~newsomj/uvclass/>

the latent variable is a math ability factor and item Y_1 is a correct/incorrect response to a problem on a standardized test, a path from the variable X could investigate bias by testing whether girls do more poorly on the item relative to boys once the loading for underlying math ability is taken into account.



Multiple-Group Models

Multiple-group or multigroup structural equation models test separate structural models in two or more groups (Jöreskog, 1971; Sorböm, 1974). Such models may involve path models, comparison of indirect effects, confirmatory factor models, or full structural equation models. Multigroup models generally follow the same structure in each group and can provide separate estimates of within-group parameters (e.g., loadings, paths, and correlations). Chi-square and fit indices can be obtained for each group separately as well as global fit indices for the joint, multigroup model. Software programs allow the user to set any number of parameters to be equal across groups, so that a single estimate of a predictive path for all of the groups, for example, is obtained. Fit for the overall multigroup model can be computed and then constraints can be imposed in a subsequent model that sets any parameter or set of parameters (e.g., predictive paths, loadings, correlations, measurement residuals) equal across groups to evaluate whether there is a significant increase in chi-square. The change in chi-square is a likelihood ratio test the individual or set of parameters that were constrained to be equal. With more than two groups, one can obtain statistical tests that compare multiple groups simultaneously in an omnibus test or pairs of groups.

It is often the objective of the analysis to investigate whether a scale or test has equivalent measurement properties across groups, also referred to as invariance testing. A test that has the same measurement properties across groups is considered “invariant.”² For example, in cross-cultural research, it may be of interest to determine whether a depression scale performs equal as well in Spanish as in English. The factor structure or the loadings can be investigated across groups to determine whether underlying constructs differ or particular items perform poorly in one group or another. Establishing whether a measure has equivalent properties across groups also is of importance as an initial step in multigroup predictive analyses to ensure that the substantive group differences are not confounded with group differences in measurement properties. Although multigroup measurement analysis can be extremely valuable, the processes and details can also become exceedingly complex (Millsap, 2012). See the subsequent handout from this class called “Invariance Tests in Multigroup SEM” for more information and references.

Moderation with Continuous Variables

One approach to statistical interactions, or moderation, in SEM follows the regression approach to interactions with continuous variables. Either continuous or binary variables can be used in this approach to testing interactions, of course, but when the moderator is continuous, the multigroup approach described below cannot be used. The definition of a statistical interaction is that the effect of the predictor, X , is the same for all values of the moderator, Z . There are two general ways to represent this in path models.



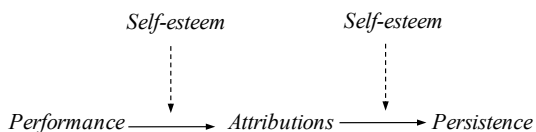
² This terminology can become a headache, because authors often use the term “noninvariant” to state that there are differences across groups.

The picture on the left is more conceptual, emphasizing that the X - Y relationship is impacted by Z . The figure on the right better reflects how the analyses are conducted, with a product variable created by multiplying x by z and then regressing Y on all three variables— x , z , and xz . The correlations among the three predictors are included because each path needs to be a partial regression with respect to the other predictors, so that the xz effect on Y represents the interaction contribution above and beyond the main effects. I substituted lower case x and z to indicate that centering is recommended for the x and z variables prior to creating the product, xz , before use in the analysis. Centering, subtracting the mean (e.g., $x = X - \bar{X}$) is recommend to reduce nonessential collinearity (Aiken & West, 1991), which avoids inflation of standard errors that affect the significance tests for the main effects of x and z .

Following significant interactions, it is also possible to conduct simple slope tests within an SEM package using model constraints to compute the simple effects and request their statistical tests (see Muthén, Muthén, & Asparouhov, 2016 and instructions and examples by Chris Stride and colleagues at <http://offbeat.group.shef.ac.uk/FIO/mplusmedmod.htm>). See also the subsequent handout “Simple Slopes for Exploring a Significant Interaction in SEM.”

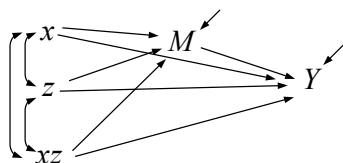
Moderated Mediation with Continuous Moderators

The moderation tests for continuous interactions can be combined with mediation analysis to investigate whether a moderator, z , moderates the relationship between a predictor x and the mediator, m , or the relationship between the mediator, m , and the outcome Y , or both. Moderated mediation is the term used to describe either of these hypothesized models (although the terms mediated moderation or conditional indirect effects might be used).³ SEM can be used to test any of the forms of moderated mediation traditionally tested with several separate regression steps in regression (e.g, Hayes, 2018; Preacher, Rucker, & Hayes, 2007), or macros that perform the separate steps automatically. If the mediational hypothesis is that self-critical attributions (m) mediate the relationship between poor performance on a test (x) and then subsequent effort persistence on the next test (Y), then self-esteem (z) could mediate either of these links in the hypothesized causal chain (James & Brett, 1984).



When the moderator is continuous, appropriate product terms, either xz or mz , must be computed prior to the analysis. Centering is again recommended for both variables.⁴ There are several different forms the model can take (Muthén et al., 2016). For tests of the indirect effects, which can be tested with a bootstrap sampling approach (see the “Testing Mediation with Regression” handout for this class), it is important to include the direct effect (i.e., the c' path).

Z moderates the x-M relationship

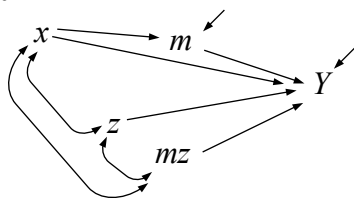


³ The terms mediated moderation and moderated mediation are a little difficult to distinguish. Mediated moderation appears to refer to an interaction between X and Z on Y that is mediated by m , whereas moderated mediation refers to the entire process or any of the process of the mediational pathway $X \rightarrow M \rightarrow Y$ that differs for different levels of Z (see MacKinnon, 2008; Muller, Judd, & Yzerbyt, 2005).

⁴ Centering is not usually done for the dependent variable, but it is inconsequential for the regression coefficient, or the test of the interaction, so, even though m may be the outcome for one of these paths, it will not hurt to use the centered version.

Z moderates the m-Y relationship

The covariances of variables mz and z with the disturbance of m could be included (Preacher et al, 2007) but they would not impact the results (Muthén et al., 2016).



Moderation using Multigroup Analysis

When the moderator is a categorical variable, predictive models, including mediational models can be used to investigate interactions. The groups should generally be naturally defined. I strongly recommend against artificial dichotomization, such as median splits, in order to form discrete groups and take advantage of the convenience of multigroup SEM. Artificial dichotomization can result a number of problems, including loss of information, loss of power, and failure to identify nonlinear effects (MacCallum Zhang, Preacher, & Rucker, 2002). Take a simple case in which the regression of Y on X is hypothesized to differ for values of Z . When Z is categorical, a multigroup SEM can be tested estimating separate slopes, $\beta_{yx|0}$ and $\beta_{yx|1}$, in one model. A second model, would then constrain the two slopes to be equal to one another, resulting in a single estimate β_{yx} in the two groups. The overall chi-square for two models can be compared with a single-*df* chi-square test to assess whether the slopes in the two groups differ significantly. The test is asymptotically equivalent to a regression in which Y is regressed on the product variable XZ .⁵

$$\begin{matrix} Z=0 \\ X \xrightarrow{\beta_{yx|0}} Y \end{matrix}$$

$$\begin{matrix} Z=1 \\ X \xrightarrow{\beta_{yx|1}} Y \end{matrix}$$

This same multigroup strategy could also be used for a latent predictor or latent outcome, or both. Use of latent variables usually presupposes the measurement invariance tests briefly described above before proceeding to the predictive analyses.

Moderated mediation hypotheses could also be investigated with the multigroup SEM approach if the moderator is categorical. For example, the mediational model, $X \rightarrow M \rightarrow Y$ could be tested in separate groups, where $Z = 0$ and $Z = 1$. Equality constraints could then be imposed on either the $X \rightarrow M$ or the $M \rightarrow Y$ path, or both paths simultaneously. The chi-square for the equality constrained model would then be compared to the chi-square for the model in which the path(s) were estimated separately and freely in the two Z groups. The difference in chi-square is a likelihood ratio test of nested models that tests whether the two groups differ significantly on whichever paths were constrained. Like the regression example above, the moderated mediation tests could involve latent variables.

Latent Variable Interactions

Latent variable interactions (Kenny & Judd, 1984) are also possible (e.g., where $\xi_1 \xi_2$ predicts Y or η), although there has been considerable exploration of how to best test them (see Kelava & Brandt, 2023, and Marsh, Wen, Nagengast, & Hau, 2012, for reviews). Latent interaction approaches generally assume normally distributed latent variables, which is one reason why a number of different approaches have been proposed. The lack of clear consensus and the fact that latent variable interaction approaches also have not been widely automated in software programs and sometimes require several steps for the user (e.g., computing product variables among indicators, complex equality constraints) has led them to be

⁵ There are some subtle (or sometimes, not so subtle) distinctions between these tests. The product regression approach assumes both the variances of X and the disturbance are equal in the two groups (the latter should sound familiar to you as the heterogeneity of variance assumption). In SEM, these assumptions could be tested through the same equality constraint process.

rarely implemented in practice. One of the approaches, the full information maximum likelihood approach (Klein & Moosbrugger, 2000; often referred to as LMS in the literature), is implemented in Mplus (using the XWITH command), however. For the R package, see Schoemann and Jorgensen (2021) for product indicator approaches in semTools. The method appears to produce unbiased parameter estimates, correct Type I error rates, and have high statistical power under at least some conditions (Cham, West, & Aiken, 2012; Cham, Reshetnyak, Rosenfeld, & Breitbart, 2017).

Extensions

I have tried to keep it simple, but any of the above analyses can be extended to more than two groups. With the product variable method, a set of dummies must be used for a categorical predictor. The process is much easier if there are more than two groups in the multigroup modeling approach, because additional groups can be declared in the analyses, with options of placing equality constraints across any number of the groups. Three-way or higher interactions could also be tested using either of the two approaches described above. There is also no reason why the two approaches cannot be combined in some way (e.g., the continuous interaction between x and z differs across W groups). I also have assumed continuous mediators and outcomes, but binary or ordinal mediators or outcomes (as well as count or multicategory variables in Mplus) could be used.

References

- Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- Cham, H., Reshetnyak, E., Rosenfeld, B., & Breitbart, W. (2017). Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behavioral Research*, *52*(1), 12-30.
- Cham, H., West, S. G., Ma, Y., & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research*, *47*(6), 840-876.
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, second edition. Guilford publications.
- James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, *69*, 307-321.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*(4), 409-426.
- Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*(351a), 631-639.
- Kelava, A., & Holger, B. (2023). Latent interaction effects. In R.H. Hoyle, *Handbook of structural equation modeling, second edition* (pp. 427-446). Guilford.
- Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological bulletin*, *96*(1), 201-210.
- Lee, J., Little, T. D., & Preacher, K. J. (2012). Methodological issues in using structural equation models for testing differential item functioning. In E. Davidov, P. Schmidt, & J. Billiet, *Cross-Cultural Analysis: Methods and Applications* (pp. 77-106). Routledge.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19-40.
- MacKinnon, D. (2008). *Introduction to statistical mediation analysis*. Routledge.
- Marsh, H.W., Wen, Z., Nagengast, B., & Hau, K-T. (2012). Structural equation models of latent interaction. In R.H. Hoyle, *Handbook of structural equation modeling* (pp. 436-458). Guilford.
- Millsap, R. E. (2012). *Statistical approaches to measurement invariance*. Routledge.
- 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.
- Muthén, B., Muthén, L.K., & Asparouhov, T. (2016). *Regression and mediational analysis using Mplus*. Muthén & Muthén.
- Newsom, J.T. (2017). Structural models for binary repeated measures: Linking modern longitudinal structural equation models to conventional categorical data analysis for matched pairs. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*, 626-635.
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*(1), 185-227.
- Schoemann, A. M., & Jorgensen, T. D. (2021). Testing and interpreting latent variable interactions using the semTools package. *Psych*, *3*(3), 322-335.
- Sorbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology*, *27*, 229-239.