Substantive Interpretation of Moderation Effects in Multilevel Logistic Regression Models

Julie Lorah
Indiana University, Bloomington
jlorah@iu.edu

Context

Many relevant educational outcomes are binary, such as graduating high school or being suspended from school. Further, educational data is often nested, such as when students are nested within schools, and ignoring this nesting can inflate Type I error rates (Snijders & Bosker, 2012). Lastly, many relationships in education are not simply linear and moderation models can account for the complexity present when the relationship between two variables depends on the level of the third (Aiken & West, 1991). To investigate data with these properties, educational researchers can use the multilevel logistic regression model:

$$\text{logit}(p_{ij}) = \beta_0 + \beta_1 {*} X_{ij} + \beta_2 {*} M_{ij} + \beta_3 {*} X_{ij} {*} M_{ij} + u_{oj} \qquad (1)$$

$$\text{Where } logit(p) = \ln(\frac{p}{1-p})$$

Where i represents the individual level (level-one) unit; j represents the group level (level-two) unit; p is the probability of a successful outcome (i.e. Y=1); β_0 is the intercept, all other β are slope coefficients, X is the independent variable, M is the moderator variable, X*M is the product term, and u_{0j} is the level-two random effect distributed with mean of zero and estimated variance τ^2. The significance of the moderation effect can be assessed with a t-test on β_3 (Aiken & West, 1991).

This model can be used with experimental data, correlational data, or even to investigate psychometric questions related to differential item functioning (Gomez-Benito, Hidalgo, & Padilla, 2009). Interpretation is an important step in data analysis; however, interpretation of these models may be difficult, particularly combining the non-linear logistic regression model with a non-linear product term.

It is unclear what measure of effect size would be appropriate for the moderation effect in this model. For one-level linear regression, the f^2 measure is suggested (Aiken & West, 1991). Researchers have suggested multiple measures of variance explained for logistic regression models, although an optimal measure for interaction effects requires additional research (Hidalgo, Gomez-Benito, & Zumbo, 2014). Further, the measure for variance explained with multilevel models is unclear since variance is partitioned among two or more levels (Snijders & Bosker, 2012). The present study explores interpretation of the moderation effect in a multilevel logistic regression model by suggesting a measure of effect size for the moderation effect and providing a procedure for plotting significant moderation effects.

Research Questions

How can applied researchers interpret significant moderation effects in a multilevel logistic regression model? Specifically:

1. What would be an appropriate effect size measure?
2. How can this effect be plotted to interpret the result?

Methods

A literature review was conducted to identify possible effect size measures and procedures for plotting significant interaction effects. Although it's beyond the scope of this short proposal, the full paper additionally conducts a simulation study to compare multiple additional possible measures of effect size.

Additionally, these results are demonstrated for applied researchers with hypothetical data. The decision to use hypothetical data rather than real data was made since this same framework can then be utilized in the full paper which will include a simulation study. Additionally, the demonstration is not intended to provide any sort of substantive conclusions and so simulated data precludes any attempt to do so.

Data is simulated according to Equation 1 with the following properties: total sample size = 20,000; number of groups (level-2 units) = 10; $\beta_0 = \beta_1 = \beta_2 = .3$; $\beta_3 = .2$; $\tau^2 = .25$; X is normally distributed with M = 0 and SD = 1; and M is binary with p = .5. Data was generated using R (R Core Team, 2018) and models were estimated with the glmer function from the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). All syntax is included in the full paper.

Results and Demonstration

Based on a literature review, McFadden's (1974) R^2 was selected, defined as:

$$R^2 = 1 - \frac{\ln(L_M)}{\ln(L_0)} \tag{2}$$

Where ln indicates the natural log; L_M is the value of the likelihood function for the model being estimated and L_0 is the value of the likelihood function for a model with no predictors. Based on likelihood values rather than variance component estimates, this measure represents a general measure applicable when maximum likelihood estimation is used.

The effect size measure for a moderation effect (Aiken & West, 1991) is f^2 computed as:

$$f^2 = \frac{R_2^2 - R_1^2}{1 - R_2^2} \tag{3}$$

Where R^2_1 is based on the main effects only model (i.e. formula 1 but omitting the product term) and R^2_2 indicates the full interaction model (i.e. formula 1). This measure can be interpreted as the proportion of variance accounted for by the interaction effect relative to unexplained variance (Aiken & West, 1991). For the present study, McFadden's (1974) R^2 is used for these computations.

For the demonstration analysis, R^2_1 (main effects only model) was .0287; R^2_2 (interaction model) was .0298 and f^2 was .0011.

The simple slopes procedure described by Aiken & West (1991) is generalized to the multilevel logistic regression model. Specifically, the predicted probabilities (probability Y=1) can be computed (Snijders & Bosker, 2012) according to:

$$p = \frac{e^{\beta0 + \beta1*Xij + \beta2*Mij + \beta3*Xij*Mij + uoj}}{1 + e^{\beta0 + \beta1*Xij + \beta2*Mij + \beta3*Xij*Mij + uoj}} \tag{4}$$

Where variables are defined according to those in formula 1 (Snijders & Bosker, 2012). Curves representing the relationship between X and Prob(Y=1) are plotted for various levels of M and U_{oj} using the *curve* function in R (see figures 1 & 2).

Conclusions

This proposal has defined and demonstrated important aspects related to interpretation of the multilevel logistic regression moderation model, including suggesting a measure of effect size and providing and demonstrating a procedure for plotting these effects. Due to space constraints, additional topics not included here will be included in the full paper including generalization and demonstration of findings from level-one interaction terms to level-two interactions, cross-level interactions, and random slopes (fixed & random component interactions). Further, a simulation study is conducted to compare the proposed effect size measure to other measures suggested in the literature and an appendix with all R syntax used to simulate data and conduct analyses is provided.

References

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Bates, D., Maechler, M., Bolker,B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01.

Gomez-Benito, J., Hidalgo, M. D., Padilla, J. (2009). Efficacy of effect size measures in logistic regression: An application for detecting DIF. *Methodology: European Journal of Research Methods for the Behavioral & Social Sciences*, *5*(1), 18-25.

Hidalgo, M., Gomez-Benito, J. & Zumbo, B. D. (2014). Binary logistic regression analysis for detecting differential item functioning: Effectiveness of R^2 and delta log odds ratio effect size measures. *Educational and Psychological Measurement, 74*(6), 927-949.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105-142). Academic Press.

R Core Team (2018). R: A language and environment for statistical computing [computer software]. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Los Angeles: Sage.

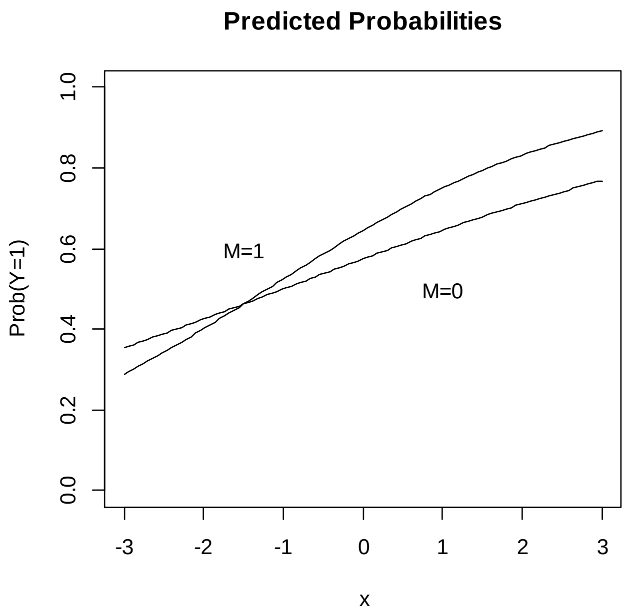

Figure 1

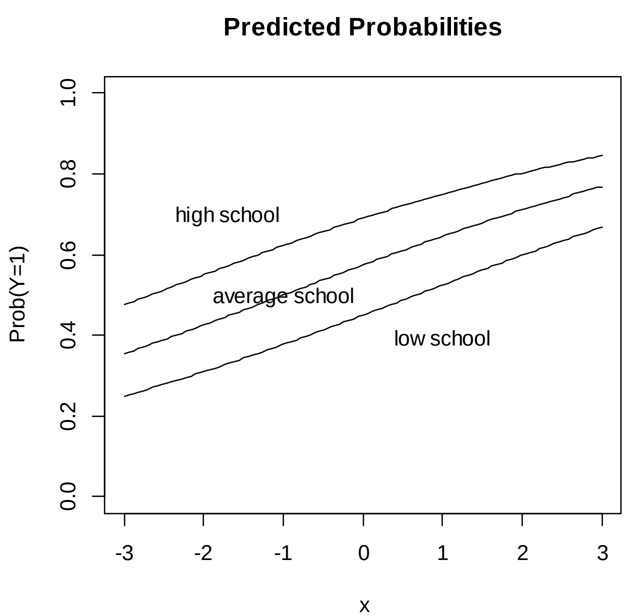

Figure 2