[Journal of Modern Applied Statistical Methods](#)

Volume 13 | Issue 1 Article 2

5-1-2014

Relative Importance of Predictors in Multilevel Modeling

Yan Liu
Harvard University, yan_liu@hms.harvard.edu

Bruno D. Zumbo
University of British Columbia, bruno.zumbo@ubc.ca

Amery D. Wu
University of British Columbia, ameryw@yahoo.com

Journal of Modern Applied Statistical Methods
May 2014, Vol. 13, No. 1, 2-22.

ISSN 1538 − 9472

Regular Articles: Relative Importance of Predictors in Multilevel Modeling

Yan Liu
Harvard University
Cambridge, MA

Bruno D. Zumbo
Univ. of British Columbia
Vancouver, BC, CAN

Amery D. Wu
Univ. of British Columbia
Vancouver, BC, CAN

The Pratt index is a useful and practical strategy for day-to-day researchers when ordering predictors in a multiple regression analysis. The purposes of this study are to introduce and demonstrate the use of the Pratt index to assess the relative importance of predictors for a random intercept multilevel model.

Keywords: Random Intercept model, multilevel model, M*plus*, Structural equation modeling, Pratt Index

Introduction

Multiple regression analysis is a widely used statistical method in many fields. Once predictors in a regression model are selected, it is a common practice for researchers to investigate which predictors explains more variance than others, or to identify a sub-set of predictors that explain most of the variation in the outcome variable. Hence, how to measure the relative importance of explanatory variables has been widely discussed in the regression literature (e.g., Budescu, 1993; Darlington, 1968; Green, Carroll, & Desarbo, 1978; Kruskall, 1987; Pratt, 1987; Thomas, Hughes, & Zumbo, 1998). As is commonly noted in the literature, the relative importance of a predictor reflects how much it contributes to the explanation/prediction of an outcome variable, in the presence of the other correlated predictors.

The Pratt index, a R-square based statistic, has been shown to be a useful and practical strategy when ordering predictors in terms of importance in a

Dr. Liu has a Ph.D. in Psychometrics. Email at: yan_liu@hms.harvard.edu. Dr. Zumbo is Professor of Measurement and Associate Editor of JMASM. Email him at: bruno.zumbo@ubc.ca. Dr. Wu is Assistant Professor of Measurement. Email her at: ameryw@yahoo.com.

multiple regression analysis. However, to date, this technique has not been adapted for multilevel model or hierarchical linear model (HLM) analysis because (a) there is no natural R-square measure for a multilevel regression model akin to one in multiple regression that can be partitioned additively, and (b) the within- and between-level correlation matrices are not readily available – both of which are key elements in R-square based methods for variable ordering. However, recent advances on multilevel modeling within a structural equation modeling (SEM) framework provides these two key elements (e.g., Asparouhov & Muthén, 2006; Muthén, 1994) and hence allows one to apply the Pratt index to multilevel regression models.

As Raudenbush and Bryk (2002) note, the random intercept model is widely used, especially when the clustering is a nuisance factor or one is interested in how the level-2 predictors affect the means of the outcome variable (e.g., Bryk & Driscoll, 1988; Englert, et al., 1988; Judge, Scott, & Ilies, 2006; Muijs, 2003), and hence the ordering of predictors has practical significance and value. The purpose of this study is to demonstrate how to order the relative importance of predictors in a multilevel regression analysis with a random intercept using the Pratt index (Pratt, 1987; Thomas, Hughes, & Zumbo, 1998; Zumbo, 2007). The article is organized as follows. First, the Pratt index is briefly described. "Next, the additive property of R-square measures and estimated covariance matrices at within- and between-levels are described. Finally, it will be demonstrated how to use the Pratt index in multilevel regression models using M*plus* with two examples: (a) a random intercept only multilevel regression analysis, and (b) a random intercept only with a new multilevel regression approach-- latent covariate.

Pratt Index

Herein a very brief sketch of Pratt's variable ordering measure is provided, similar to the one described in Zumbo (2007). The interested reader is referred to Pratt (1987) and Thomas, Hughes, & Zumbo (1998) for details. Pratt considered a linear regression of the form

$$y = b_0 + b_1 x_1 + \ldots + b_p x_p + \varepsilon \,, \tag{1}$$

where residual term ε is uncorrelated with $x_1, x_2, \ldots, x_p$ and is distributed with mean zero and variance σ^2. The total standardized variance (R^2) in a population explained by the model in equation (1) can be written as

$$R^2 = \sum_j b_j \rho_j \tag{2}$$

where b_j is the standardized regression coefficient corresponding to x_j, and ρ_j is the simple correlation (i.e., zero-order correlation) between y and x_j. Pratt justified the rule whereby relative importance of a predictor is equated to variance explained, provided that the explained variance attributed to x_j is $b_j \rho_j$, a definition which is widely used in the applied literature (e.g., Green, Carroll & De Sarbo, 1978).

An additional feature of Pratt's measure is that it allows the importance of a subset of variables to be defined additively, as the sum of their individual importance irrespective of the correlation among the predictors. Other commonly used measures (e.g., the standardized beta-weights, the *t*-values, zero-order correlations, semi-partial correlations) do not allow for an additive definition and may be problematic with correlated predictor variables.

Thomas, Hughes, and Zumbo (1998) provide a sample interpretation of Pratt's measure based on the geometry of least squares. They considered a sample regression equation,

$$\hat{y} = \hat{b}_1 x_1 + \ldots + \hat{b}_p x_p \tag{3}$$

where the $\hat{b}_j$ s are estimates of the population regression coefficients, $j = 1, \ldots, p$. They defined the partition of R^2 of x_j, j=1, ..., p, to be the signed length of the orthogonal projection of $\hat{b}_j x_j$ onto $\hat{y}$, to the length of $\hat{y}$. By definition, this ratio represents the proportion of R^2 and sums to 1.0. Furthermore, the partitioning is additive, so that one could, for example, compute the proportion of R^2 attributable to various subsets of the explanatory variables, irrespective of the correlations among the explanatory variables.

One then can partition the resulting R^2 by computing the Pratt index, d_j,

$$d_j = \frac{\hat{b}_j \times r_j}{R^2}, \tag{4}$$

where, as above, $\hat{b}_j$ is the *j*th standardized regression coefficient (the "beta"), r_j is the simple Pearson correlation, also called zero-order correlation, between the response variable and *j*th explanatory variable in equations (1) and (3) in samples.

The sum of d_j, computed from equation (4), over all predictors is one, and the relative importance of predictors can be ordered by d_j, that is, the larger the value of d_j the more important the predictor, as per Pratt (1987). Thomas (1992) suggested that as a general rule, if $d_j < 1/(2p)$ (where p is the number of predictors), namely half the average importance, then the predictor can be regarded as unimportant.

A variety of strategies have been used in practice in the literature, such as standardized regression coefficients (i.e., beta-weights), zero-order correlations, and the *t*-tests and its *p*-values for the regression coefficients, but they can give inconsistent results when the predictors are correlated because they do not have the additive property as indicated above. In addition to the Pratt index, two other methods have also been recommended in the literature, dominance analysis (Budescu, 1993) and proportional marginal variance decomposition (i.e., a modified version of dominance analysis) (Feldman, 2005). However, these two methods are computationally intensive with even a modest number of predictors, whereas the Pratt index requires simple computation and is easy to understand and interpret.

Additive R-squares and Correlations Using SEM

R-square is a widely used global effect size in multiple regression analysis, which is used to quantify the variance in an outcome variable explained by the model (i.e., by all the explanatory variables). However, R-square in a multilevel analysis is not straightforward. Several R-square or effect size measures were suggested in the literature, but none of them is equivalent to the one used in a multiple regression and the calculation of R-square for a random slope model is more complex due to the covariance of residuals between the intercept and slope(s) (Gelman & Pardoe, 2006; Hox, 2010; Kreft & de Leeuw, 1998; Raudenbush & Bryk, 2002; Roberts & Monaco, 2006; Singer & Willett, 2003; Snijders & Bosker, 1999).

Based on a SEM framework, the recent advances in multilevel modeling have made it possible for us to use the Pratt index in a multilevel regression analysis. Unlike the conventional multilevel modeling approach, the observed covariance matrix can be decomposed into within- and between-levels orthogonally using the SEM framework. Cronbach and Webb (1979) proposed to decompose the observed individual variables into within- and between- group components, which can be written as $Y_{tot} = Y_w + Y_b$, and the components Y_w and Y_b are orthogonal and additive. This decomposition can be used for the partition of population covariance matrix to Σ_w (within-level covariance matrix) and Σ_b (between-level covariance matrix). Muthén (1989, 1990, 1994) showed that the sample covariance matrices can be used to estimate the multilevel population covariance matrices. In addition, Muthén (1994) showed that the pooled within-level covariance matrix is an unbiased estimate of the population within-level covariance matrix Σ_w, which is given by

$$S_w = \frac{\sum_j^G \sum_i^n (Y_{ij} - \bar{Y}_j)(Y_{ij} - \bar{Y}_j)'}{N - G} = \hat{\Sigma}_w$$

where i denotes individuals, j denotes groups, N is the total sample size, G is the total number of groups, Y_{ij} denotes individual observations of all observed variables, $\bar{Y}_j$ denotes the group means of all observed variables, and the symbol prime denotes transpose. Muthén further showed that the sample between-level covariance matrix is an estimate of the composite $S_b = \hat{\Sigma}_w + c\hat{\Sigma}_b$, where c is a scaling factor

$$c = \frac{N^2 - \sum_j^G n_j^2}{N(G-1)}$$

and S_b is given by

$$S_b = \frac{\sum_j^G n(\bar{Y} - \bar{Y}_j)(\bar{Y} - \bar{Y}_j)'}{G - 1}.$$

The maximum likelihood estimate of $\hat{\Sigma}_b$ is $c^{-1}(S_b - S_w)$.

The estimated within- and between-level covariance matrices allow us to obtain two key components that are needed for the calculation of the Pratt index: the correlations of the outcome variable with the predictors and the variances of the outcome variable explained by the model at within- and between-levels, respectively. The correlations can be obtained from covariance matrices as the correlation matrices are simply the standardized covariance matrices. The additive property of estimated variance-covariance matrices at the within- and between-levels makes it possible to obtain the R-square which is conceptually equivalent to the one used in a multiple regression analysis and is always positive – a property that is not always guaranteed by other "R-square" measures discussed in the literature. The total variance of the outcome variable at both levels can be obtained directly from the covariance matrices, the residual variances at both levels can be obtained from a multilevel model analysis, and R-square can be computed from the equation $R^2 = 1 - \frac{\sigma_e^2}{\sigma_{tot}^2}$, which applies to both within- and between-levels. It should be noted that R-square arising from this method is akin to the R-square in regression and hence can be partitioned using Equation (4). However, it should be noted that the additive property of R-square described here only applies to a random intercept regression model and the problems raised by a random slope model are still not yet solved. This limitation is also true for other methods of ordering in multilevel models such as those based on dominance analysis (Luo & Azen, 2013).

The M*plus* software program has currently made those parameter estimates available in the output file. The covariance as well as correlation matrices at both within- and between-levels can be obtained by requesting "SAMPSTAT" under the "OUTPUT" command. The request for "STANDARDIZED" under the "OUTPUT" command will give R-squares for within- and between-levels, respectively, and the standardized beta-weights (i.e., beta-weights in the section of "STDYX Standardization" in the output). Researchers can also calculate R-square using the variance of outcome variable and the residual variance obtained from the M*plus* output. Examples of M*plus* syntax and M*plus* output can be found in Appendices A and B.

Two Demonstrations

In this section, the use of the Pratt index with two real data examples is demonstrated. The first is a demonstration of a commonly used model in conventional HLM practice and involves what is often referred to as a random intercept model with predictors at both within- and between-levels. The second is a demonstration of a model that is referred to as a latent covariate approach, wherein the observed predictors are decomposed into two latent components rather than the common practice of aggregating individual observations to form a group level predictor.

Data sources

The data were retrieved from Trends in the International Mathematics and Science Study (*TIMSS*) 2007. TIMSS 2007 Grade-8 mathematics ability scores, plausible values, estimated by item response theory, were used as the outcome variable. For the purpose of demonstration, one of five plausible values for the analyses was chosen. Six predictors (either measured variables or derived indices by TIMSS) were chosen from the students' questionnaire as within-level predictors. These within-level predictors included sex, use of calculator (Calculator), availability of computer (Computer), students' positive affect toward mathematics (Affect), students' valuing of mathematics (Valuing), and students' perception about being safe at school (Safety). Three variables were chosen from the school principal's questionnaire as between-level predictors — good school attendance (Attendance), principals' perception of school climate (Climate), and percentage of students at economic disadvantage in the school (SES).

Among those predictors, Calculator and Computer are on 4-point Likert scale (never, some lessons, half the lessons, & every or almost every lesson); Affect, Valuing, Safety, Attendance, and Climate are on 3-point Likert scale (low, medium, & high); and SES are on 4-point Likert scale (0-10%, 11-25%, 26-50%, & more than 50%). A detailed description of these variables can be found in the TIMSS 2007 User Guide (Foy & Olson, 2009). A total of 120 schools, 3470 students from Hong Kong were included in the analysis and 50.4% of students are girls. It should be noted that the same data set will be used for both demonstrations.

Demonstration One

Data analysis. Please see Raudenbush and Bryk's (2002) case study one for a description of the random intercept model in their notation. In the case herein, the random intercept multilevel regression model was estimated using M*plus* 6.02 to address how students' mathematics ability was affected by the between-level (school level) factors as well as the within-level (student level) factors. The Pratt indices were computed to answer the question—which predictors are more important when accounting for the variance in the outcome variable (mathematics ability). The two-level random intercept model in the M*plus* formulation can be described as:

$$\text{Within-model:} \quad Y_{ij} = \beta_{0j} + \beta_1 \text{ Gender} + \beta_2 \text{ Valuing} + \beta_3 \text{ Computer} + \beta_4 \text{ Affect} + \beta_5 \text{ Calculator} + r_{ij}, \tag{5}$$
$$\text{Between-model:} \quad \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ Attendance} + \gamma_{02} \text{ Climate} + \gamma_{03} \text{ SES} + u_{0j},$$

where i denotes the number of students; j denotes the number of schools; β_{0j} is the random intercept; other β s are the fixed slopes of within-model predictors; γ_{00} is the model grand mean; other γ s are the slopes for the between model predictors; r_{ij} is the within-level residual; and u_{0j} is the between-level residual for the random intercept.

Results Table 1 shows the results of the multilevel regression analysis with a random intercept and the Pratt index for each predictor. The second column is the standardized regression coefficients (the 'beta'-weights); the following columns present *t*-tests, the corresponding *p*-values, zero-order correlations, and Pratt indices. The upper and lower parts of the table contain the results of the student-level (within) and school-level (between) models, respectively. The M*plus* syntax and output of Demonstrate One can be found in Appendix A.

Using the common practices described above, would make contradictory conclusions about the relative importance of within-level predictors, *sex*, *calculator*, *computer* and *valuing math*, if using different strategies. For example, one would consider *computer* more important than *sex*, *calculator* and *valuing math* if relying on the beta-weights. However, one would regard *sex* more important if relying on *t*-tests or the corresponding *p*-values or regard *valuing math* more important than computer, sex, or calculator if using simple correlations. These strategies are problematic as they do not have the additive property mentioned earlier.

However, due to its additive property, the Pratt index orthogonally partitions the R-square and sums to one, which can provide us a criterion of how much each predictor contributes to the explained variance in the outcome variable orthogonally. Using the Pratt indices, *calculator* is shown to be more important than *sex*, *computer,* and *valuing math*, but all of them have made trivial contributions to the model relative to *affect*, which accounted for 73.8% of the R-square (R^2=0.135).

Table 1. A Random Intercept Multilevel Regression Analysis and Corresponding Pratt Indices

Within Level	**Beta-weight**	**t-test**	***p*-value**	**Correlation**	**Pratt**
valuing math	0.048	2.610	0.009	0.158	0.056
computer	-0.104	-5.031	<0.001	-0.077	0.059
sex	0.090	5.108	<0.001	0.094	0.063
calculator	0.090	4.659	<0.001	0.121	0.081
Affect	0.300	16.381	<0.001	0.332	0.738
R-square	0.135				SUM=1.0
Between Level	**Beta-weight**	**t-test**	***p*-value**	**Correlation**	**Pratt**
school climate	0.249	3.831	<0.001	0.354	0.258
low SES	-0.259	-3.258	0.001	-0.430	0.326
school attendance	0.319	4.196	<0.001	0.447	0.417
R-square	0.342				SUM=1.0

Note. The sum of Pratt index of all predictors in either within- or between-levels is not exactly one due to rounding errors from parameter estimates.

For the between-level model, *school attendance* was shown as the most important predictor among the three school variables. The order of importance would also be different, depending on whether beta-weights, correlation, or *t*-tests are used as criterion. The Pratt indices showed that *school attendance* is the most important predictor, which accounted for 41.7% of the explained variance (R^2=0.342). The next important predictors are *low SES and school climate*, which accounted for 32.6% and 25.8% of the explained variance, R-square, respectively. Using Thomas' (1992) criterion, all the values of Pratt indices are greater than 0.167, so that those between-level predictors could be considered as important predictors.

Demonstration 2

Data Analysis In the second example, a new approach is demonstrated that allows us to examine a predictor at both levels though it is collected at the individual level. In some situations, data was collected from individuals, but it was also desired to investigate them at an aggregate level. For example, imagine students' socioeconomic status (SES) was collected from individual students, but also were interested in the effects of school SES. Rather than aggregating SES by taking an average from within-level, a new approach that decomposes SES variable into two latent components (SES_{within} and $SES_{between}$) in the multilevel regression analysis can be used, which would reduce the measurement error arising from aggregating the data as is typical within SEM.

In general, a manifest covariate X_{ij} can be decomposed into two latent components $X_{ij} = X_{wij} + X_{bj}$ where X_{wij} and X_{bj} are latent covariates. The multilevel equations are defined as $Y_{ij} = \beta_{0j} + \beta_{1j} X_{wij} + r_{ij}$ and $\beta_{0j} = \gamma_{00} + \gamma_{01} X_{bj} + u_{0j}$ where all the notation is defined the same as in Equation (3). A detailed description can be found in Asparouhov and Muthen (2006) and Ludtke, et al. (2008). This approach has also been adopted in Preacher, Zyphur, and Zhang's (2010) multilevel mediational models.

This latent variable decomposition approach is used in the second demonstration. Building on the model in the first demonstration, one variable *safety* (How safe students feel at schools) was added, which was collected from individual students, but the effects of safety at both student and school levels can be examined using this latent variable decomposition approach. The 2-level model is presented as follows:

$$\begin{aligned}
&\text{Within-model:} \quad Y_{ij} = \beta_{0j} + \beta_1 \text{ Gender} + \beta_2 \text{ Valuing} + \beta_3 \text{ Computer} \\
&\qquad\qquad + \beta_4 \text{ Affect} + \beta_5 \text{ Calculator} + \beta_6 \textbf{ Safety}_{\text{within}} + r_{ij}, \\
&\text{Between-model:} \quad \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ Attendance} + \gamma_{02} \text{ Climate} + \gamma_{03} \text{ SES} \\
&\qquad\qquad + \gamma_{04} \textbf{ Safety}_{\text{between}} + u_{0j},
\end{aligned} \tag{6}$$

where all parameters are defined as in the Equation (3) except that the variable *safety* was added into the within-level and between-level models in Equation (4); β_6 is the slope for *safety* in level-1 model; γ_{04} is the slope for *safety* in level-2 model.

Results Table 2 presents the parameter estimates obtained from the random intercept multilevel regression analysis based on a latent variable decomposition approach. The columns 2-5 are the standardized beta-weights, t-tests, corresponding *p*-values, and correlations, respectively. The Pratt indices are calculated and shown in the last column. The M*plus* syntax and output of Demonstrate Two can be found in Appendix B.

The interesting findings of this analysis are that although it was collected at the individual level *safety* was a trivial predictor at the student level, beta-weight=0.008, t=0.478, *p*=0.632, and the corresponding Pratt index=0.001, but it became a salient predictor at the school level, beta-weight=0.306, t=2.925, *p*=0.004, and the corresponding Pratt index=0.306.

Table 2. A Random Intercept Multilevel Regression Analysis Based on a Latent Variable Decomposition Approach and the Corresponding Pratt Indices

Within Level	Beta-weight	t-test	*p*-value	Correlation	Pratt
Safety	**0.008**	**0.478**	**0.632**	**0.005**	**0.001**
valuing math	0.048	2.657	0.008	0.158	0.056
computer	-0.102	-4.920	<0.001	-0.075	0.056
sex	0.093	5.168	<0.001	0.097	0.066
calculator	0.093	4.763	<0.001	0.124	0.085
Affect	0.300	16.365	<0.001	0.333	0.735
R-square	0.136				SUM=1.0

Between Level	Beta-weight	t-test	*p*-value	Correlation	Pratt
school climate	0.228	3.840	0.005	0.355	0.185
low SES	-0.239	-3.042	0.002	-0.430	0.235
school attendance	0.274	3.755	<0.001	0.447	0.280
Safety	**0.306**	**2.925**	**0.004**	**0.427**	**0.306**
R-square	0.436				SUM=1.0

Note. The sum of Pratt index of all predictors in either within- or between-levels is not exactly one due to rounding errors from parameter estimates.

Again, the relative importance of predictors would be ordered differently, depending on which criterion, beta-weights, t-tests, or correlations, were used for the judgment. For example, the effect of *safety* at the between level would not be considered as the most important predictor if the judgment is based on correlations or t-tests. However, using Pratt indices, *safety* was regarded as the most important predictor, which accounted for 30.6% of the R-square. The importance of the other between-level predictors is ranked in the following the order, *school attendance* (28%), *low SES* (23.5%), and *school climate* (18.5%). The order of relative importance for the within-level predictors was similar to that of Demonstration One except the inclusion of *safety*, which should not be regarded as an important predictor based on 1/(2p) criterion as it accounted for less than 8.3% of the R-square.

Concluding Remarks

Ordering the relative importance of predictors has been a common practice in multiple regression analysis, but the methods developed in multiple regression have not been used in multilevel regression analysis due to several statistical challenges. This study demonstrated how to order the relative importance of predictors in a multilevel regression analysis using the Pratt index. The Pratt index has not been used in multilevel regression analyses mainly because the within- and between-level variance-covariance could not be partitioned orthogonally and thus an R-square measure equivalent to the one used in multiple regression analysis cannot be obtained. The recent advances in multilevel modeling using SEM framework made the R-square available for researchers to compute the Pratt index when conducting a random intercept multilevel regression analysis.

The Pratt index provides a useful tool to day-to-day researchers. As indicated in the introductory section, the Pratt index can be used with random intercept models when one wants to eliminate a nuisance factor arising from clustering or when one is only interested in the relationship between the level-2 predictors and the average scores of outcome.

It should be noted that the Pratt index can currently only apply to a random intercept regression model. The problems of obtaining a R-square with a random slope described above have not been solved yet, such as the R-square values can be negative and the magnitude of global R-square measure depends on the scale of predictors included in the model. The residual covariances of the intercept and slopes give rise to the complexity of partitioning the within- and between-level variances-covariances. Luo and Azen (2013) also discussed this issue in their extension of dominance analysis to multilevel models. They pointed out that the random slope model is problematic when conducting dominance analysis and hence suggested readers to use the random intercept model when using dominance analysis.

It is worth noting that the Pratt index is not used as a strategy to select variables, but a tool for ordering the relative importance of variables once predictors have been chosen. Selection of the variables should be based on the data as well as the theories and literature surrounding the dependent and predictor variables. Moreover, Pratt index can only tell us the statistical importance of variables, but in practice researchers also need to consider the practical/substantive importance of variables.

References

Asparouhov, T., & Muthen, B. (2006). *Constructing covariates in multilevel regression (Mplus Web Notes No. 11)*. Retrieved from http://www.statmodel.com.

Bring, J. (1996). A geometric approach to compare variables in a regression model. *American Statistician, 50*, 57-62.

Bryk, A. S. , & Driscoll, M. E. (1988). *An empirical investigation of school as a community*. Madison, WI: University of Wisconsin Research Center on Effective Secondary Schools.

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin, 114*, 542 - 551.

Cronbach, L. J., & Webb, N. (1979). Between class and within class effects in a reported aptitude × treatment interaction: A reanalysis of a study by G. L. Anderson. *Journal of Educational Psychology, 67*, 717-724.

Darlinton, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*, 161-182.

Englert, C. S., Raphael, T. E., Anderson, L. M., Anthony, H. M., Fear, K. L., & Gregg, S. L. (1988). *A case for writing intervention: Strategies for writing informational text*. East Lansing, MI: Michigan State University, Institute for Research On Teaching.

Feldman, B. (2005). *Relative Importance and Value*, unpublished manuscript (Version 1.1, March 19, 2005). Available online at http://www.prismanalytics.com/ docs/ RelativeImportance050319.pdf.

Foy, P. & Olson, J.F. (Eds.). (2009). *TIMSS 2007 International database and user guide (Trends in International Mathematics and Science Study).* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Gelman, A., & Pardoe, I. (2006). Bayesian measures of explained variance and pooling in multilevel (hierarchical) models. *Technometrics, 48*, 241-251.

Green, P. E., Carroll, J. D., & Desarbo, W. S. (1978). New Measure of Predictor Variable Importance in Multiple-Regression. *Journal of Marketing Research, 15*, 356-360.

Hoffman, P. J. (1962). Assessment of the independent contributions of predictors. *Psychological Bulletin, 59*, 77-80.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Judge, T., Scott, B., & Ilies, R. (2006). Hostility, job attitudes and workplace deviance: Test of a multilevel model. *Journal of Applied Psychology, 91,* 126 – 138.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Newbury Park, CA: Sage.

Kruskal, W. (1987). Relative importance by averaging over orderings. *American Statistician, 41*, 6-10.

Ludtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*, 203–229.

Luo, W., & Azen, R. (2013). Determining predictors importance in hierarchical linear models using dominance analysis. *Journal of Educational and Behavioral Statistics, 38*, 3-31.

Muijs, D. (2003). The effectiveness of learning support assistants in improving the mathematics in primary school. *Educational Research, 45*, 219–230.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. Presidential address to the Psychometric Society, July, 1989. *Psychometrika, 54*, 557-585.

Muthén, B. (1990). Mean and covariance structure analysis of hierarchical data. Paper presented at the Psychometric Society meeting in Princeton, NJ, June 1990. *UCLA Statistics Series 62*.

Muthén, B. (1994). Multilevel covariance structure analysis. In J. Hox, & I. Kreft (Eds.), Multilevel Modeling, a special issue of *Sociological Methods & Research, 22*, 376-398.

Pratt, J. W. (1987). Dividing the indivisible: using simple symmetry to partition variance explained. In T. Pukkila and S. Puntanen (eds.), *Proceedings of the Second International Conference in Statistics* (pp. 245-260). Tampere, Finland: University of Tampere.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209-233.

Raudenbush, S. W. Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. 2nd edition. Newbury Park , CA : Sage.

Roberts, J. K., & Monaco, J. P. (April, 2006). *Effect size measures for the two- level linear multilevel model*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, UK: Oxford University Press.

Snijders, T. A. B., & Bosker, R. (1994). Modeled variance in two-level models. *Sociological Methods & Research, 22*, 324-363.

Southwell, B. (2005). Between messages and people: A multilevel model of memory for television content. *Communication Research*, *32,* 11 -140.

Thomas, D. R. (1992). Interpreting discriminant functions: A data analytic approach. *Multivariate Behavioural Research, 27*, 335-362.

Thomas, D. R., Hughes, E., & Zumbo, B. D. (1998). On variable importance in linear regression. *Social Indicators Research, 45*, 253-275.

Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C.R. Rao and S. Sinharay (Eds.) *Handbook of statistics, Vol. 26: Psychometrics,* (pp. 45-79). Elsevier Science B.V.: The Netherlands.

Appendix A: M*plus* Syntax and Output for Demonstration One

M*plus* Syntax

```
TITLE:  this is an example of a two-level regression analysis for a continuous
        dependent variable with a random intercept and observed covariates
DATA:   FILE = HK_example2.dat
        FORMAT ARE 136.F8;
VARIABLE:
    NAMES = idsch idstd y1-y5 sex calculator computer affect valuing
                      paredu safty attendan climate ses;
    USEVARIABLES = y1 sex calculator computer affect valuing
                      attendan climate ses;
    MISSING = blank;
    WITHIN = sex calculator computer affect valuing
    BETWEEN = attendan climate ses;
    CLUSTER = idsch;
    CENTERING = GRANDMEAN (sex calculat computer affect valuing);
ANALYSIS:      TYPE = TWOLEVEL;
MODEL:
    %WITHIN%
        y1 ON sex calculat computer affect valuing;
    %BETWEEN%
        y1 ON attendan climate ses;
OUTPUT:        SAMPSTAT  STANDARDIZED;
```

M*plus* Output

```
SAMPLE STATISTICS


ESTIMATED SAMPLE STATISTICS FOR WITHIN
Covariances
           Y1         SEX        CALCULAT   COMPUTER   AFFECT     VALUING
Y1         3259.376
SEX        2.687      0.250
CALCULAT   5.823      -0.037     0.706
COMPUTER   -3.661     0.051      0.030      0.700
AFFECT     16.487     0.035      0.093      0.025      0.757
```

VALUING	5.847	0.007	0.062	0.024	0.192	0.420

Correlations

	Y1	SEX	CALCULAT	COMPUTER	AFFECT	VALUING
Y1	1					
SEX	0.094	1				
CALCULAT	0.121	-0.089	1			
COMPUTER	-0.077	0.122	0.042	1		
AFFECT	0.332	0.081	0.127	0.035	1	
VALUING	0.158	0.022	0.114	0.045	0.34	1

ESTIMATED SAMPLE STATISTICS FOR BETWEEN

Covariances

	Y1	ATTENDAN	CLIMATE	SES
Y1	5091.540			
ATTENDAN	19.053	0.357		
CLIMATE	14.287	0.049	0.319	
SES	-32.546	-0.225	-0.138	1.127

Correlations

	Y1	ATTENDAN	CLIMATE	SES
Y1	1			
ATTENDAN	0.447	1		
CLIMATE	0.354	0.144	1	
SES	-0.43	-0.355	-0.23	1

STANDARADIZED MODEL RESULTS

STDYX Standardization

Within Level

	Estimate	S.E.	Est./S.E.	P-Value
Y1 ON				
SEX	0.090	0.018	5.108	0.000
CALCULAT	0.090	0.019	4.659	0.000
COMPUTER	-0.104	0.021	-5.031	0.000
AFFECT	0.300	0.018	16.381	0.000
VALUING	0.048	0.018	2.610	0.009
Residual Variances				
Y1	0.865	0.012	69.355	0.000

Between Level

```
Y1 ON
ATTENDAN   0.319      0.076      4.196      0.000
CLIMATE    0.249      0.065      3.831      0.000
SES        -0.259     0.080      -3.258     0.001
Intercepts
Y1         7.939      0.554      14.319     0.000
Residual Variances
Y1         0.658      0.077      8.557      0.0000

R-SQUARE
Within Level
Observed
Variable   Estimate   S.E.       Est./S.E.  P-Value
Y1         0.135      0.012      10.786     0.000
Between Level
Y1         0.342      0.077      4.448      0.000
```

Appendix B: M*plus* Syntax and Output for Demonstration Two

M*plus* Syntax

```
TITLE:  this is an example of a two-level regression analysis for a continuous
        dependent variable with a random intercept – latent variable
        decomposition
DATA:   FILE = HK_example2.dat
        FORMAT ARE 136.F8;
VARIABLE:
    NAMES = idsch idstd y1-y5 sex calculator computer affect valuing
                      paredu safty attendan climate ses;
    USEVARIABLES ARE y1 sex calculator computer affect valuing
                      safty attendan climate ses;
    MISSING = blank;
    WITHIN = sex calculator computer affect valuing
    BETWEEN = attendan climate ses;
    CLUSTER = idsch;
    CENTERING = GRANDMEAN (sex calculat computer affect valuing
                      safty attendan climate ses);
ANALYSIS:      TYPE = TWOLEVEL;
MODEL:
    %WITHIN%
        y1 ON sex calculat computer affect valuing safty;
    %BETWEEN%
        y1 ON attendan climate ses safty;
OUTPUT:        SAMPSTAT  STANDARDIZED;
```

M*plus* Output

```
SAMPLE STATISTICS


ESTIMATED SAMPLE STATISTICS FOR WITHIN
Covariances
          Y1        SEX       CALCULAT  COMPUTER  AFFECT    VALUING   SAFTY
Y1        3255.204
SEX       2.757     0.250
CALCULAT  5.944     -0.038    0.706
```

COMPUTER	-3.555	0.052	0.030	0.699			
AFFECT	16.522	0.036	0.093	0.025	0.757		
VALUING	5.853	0.007	0.062	0.024	0.191	0.420	
SAFTY	0.188	-0.040	-0.001	-0.024	0.009	-0.002	0.415

Correlations

	Y1	SEX	CALCULAT	COMPUTER	AFFECT	VALUING	SAFTY
Y1	1						
SEX	0.097	1					
CALCULAT	0.124	-0.091	1				
COMPUTER	-0.075	0.124	0.043	1			
AFFECT	0.333	0.084	0.127	0.034	1		
VALUING	0.158	0.023	0.114	0.044	0.339	1	
SAFTY	0.005	-0.126	-0.001	-0.045	0.016	-0.004	1

ESTIMATED SAMPLE STATISTICS FOR BETWEEN

Covariances

	Y1	SAFTY	ATTENDAN	CLIMATE	SES
Y1	5093.476				
SAFTY	4.285	0.020			
ATTENDAN	19.068	0.015	0.357		
CLIMATE	14.302	0.009	0.049	0.319	
SES	-32.546	-0.020	-0.225	-0.138	1.127

Correlations

	VALUING	SAFTY	ATTENDAN	CLIMATE	SES
VALUING	1				
SAFTY	0.427	1			
ATTENDAN	0.447	0.183	1		
CLIMATE	0.355	0.108	0.144	1	
SES	-0.430	-0.137	-0.355	-0.230	1

STANDARDIZED MODEL RESULTS

STDYX Standardization

Within Level

	Estimate	S.E.	Est./S.E.	P-Value
Y1 ON				
SEX	0.093	0.018	5.168	0.000
CALCULAT	0.093	0.020	4.763	0.000
COMPUTER	-0.102	0.021	-4.920	0.000

AFFECT	0.300	0.018	16.365	0.000
VALUING	0.048	0.018	2.657	0.008
SAFTY	0.008	0.016	0.478	0.632
Residual Variances				
Y1	0.864	0.013	67.907	0.000
Between Level				
Y1 ON				
ATTENDAN	0.274	0.073	3.755	0.000
CLIMATE	0.228	0.059	3.843	0.000
SES	-0.239	0.079	-3.042	0.002
SAFTY	0.306	0.105	2.925	0.003
Intercepts				
Y1	7.972	0.546	14.601	0.000
Residual Variances				
Y1	67	03	96	00

R-SQUARE

Within Level

Observed Variable	Estimate	S.E.	Est./S.E.	P-Value
Y1	0.136	0.013	10.660	0.000
Between Level				
Y1	0.433	0.103	4.197	0.000