DOCUMENT RESUME

ED 453 261 TM 032 798

AUTHOR Koenig, Judith A.; Lissitz, Robert W.
TITLE The Effects of Centering Method on Parameter Estimates and Variance Explained in Multilevel Models.
PUB DATE 2001-04-11
NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Estimation (Mathematics); High Schools; *Scaling
IDENTIFIERS *Centering; *Multilevel Analysis; National Education Longitudinal Study 1988; Variance (Statistical)

ABSTRACT

This paper uses data from the National Education Longitudinal Study of 1988 (NELS:88) dataset to demonstrate practical examples of the ways in which the method used for centering level-1 variables in multilevel models affects the findings. Demonstrations compare raw metric scaling, grand mean centering, and group mean centering for successively more complex models. Comparisons are made of parameter estimates, their significance levels, and increments in variance explained. Findings show that results are generally similar for raw metric scaling and grand mean centering, and these results differ from those obtained under group mean centering. Two methods are demonstrated for estimating incremental variance explained by nested models. The ways in which centering can be used to examine between-groups and within-groups effects are also shown. (Contains 6 tables, 2 figures, and 11 references.) (Author)

The Effects of Centering Method on Parameter Estimates and Variance Explained in Multilevel Models

Judith A. Koenig
Robert W. Lissitz, Ph.D.
University of Maryland

Abstract

This paper uses data from the NELS:88 dataset to demonstrate practical examples of the ways in which the method used for centering level-1 variables in multilevel models affects the findings. Demonstrations compare raw metric scaling, grand mean centering, and group mean centering for successively more complex models. Comparisons are made of parameter estimates, their significance levels, and increments in variance explained. Findings show that results are generally similar for raw metric scaling and grand mean centering, and these results differ from those obtained under group mean centering. Two methods are demonstrated for estimating incremental variance explained by nested models. The ways in which centering can be used to examine between-groups and within-groups effects are also shown.

Paper presented at the Annual Meeting of the American Educational Research Association
Seattle, Washington
April 11, 2001
Authors can be reached at jkoenig@nas.edu

Introduction

Multilevel modeling techniques provide a way to analyze nested data. Data obtained from organizational or educational settings are inherently nested, given that individuals are nested in offices or classrooms, offices or classrooms are nested in buildings or schools, and so on. Hierarchical linear modeling (HLM) is one computer program for analyzing such data. The model used by HLM, attempts to explain the effects of independent variables on some outcome variable. The level 1 model examines the relationships among predictors and outcome variables for individuals, much like ordinary least squares regression models. With HLM, however, the intercept and regression coefficients from the level 1 model, conceptually become the dependent variables in the level 2 model. In the level 2 model, group-level variables are used to explain the between-group variance in the level 1 parameters. Such models are useful in distinguishing the effects of individual-level characteristics from group-level characteristics on the outcome measure.

Centering is an important consideration in HLM. As with multiple regression, the intercept is defined as the value of the outcome variable when the predictor variable(s) is zero. For some predictor variables, values of zero are meaningless (i.e., developmental age) or out of range (i.e., SAT scores). Since HLM focuses on explaining variance in the intercept and regression coefficients, it is critical that their meanings be clear. Unlike multiple regression, the centering transformations that are routinely used can have a substantial impact on the results and the interpretation of the regression equations.

To date, some research has been conducted on the effects of centering. While prior research is invaluable, it has not provided concrete examples as well as theoretical approaches for both simple and complex models in an educational setting and in a manner understandable to newer users of HLM. The present study attempts to fill this gap by examining two questions:

1. What are the implications of centering choices in terms of reliability, variance accounted for, and statistical significance of the parameters?
2. How should centering be used to address specific research questions?

The impact of centering methods is explained mathematically and demonstrated by application to data from an educational setting using successively more complex models. The data for this paper come from the National Education Longitudinal Study of 1988 (NELS:88) sponsored by the National Center for Education Statistics. Analyses focus on models commonly seen in the literature. An attempt is made to explain the impact of centering in a manner more useful to newer HLM users.

Background on Multilevel Models

In multilevel modeling, multiple models are developed, each corresponding to a certain unit of analysis. The first stage model is typically the individual-level model. A typical individual-level model would be:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij}$$

In these models, i refer to the individual and j to the group. Y_{ij} is the outcome variable measured for individuals. β_{0j} represents the intercept for a given group. β_{1j} represents the effect of a certain independent variable on the outcome for individuals. The unique effect associated with the individual is represented by r_{ij}. The individual level model can be conceived of in the same way as a multiple regression model, with β_{1j} indicating the increment in the outcome variable associated with a unit increment in X.

In the second stage model, the coefficients from the first stage model become the dependent variables. The second stage model allows the researcher to study the effects of group level variables on the variance among the values of the coefficients. A typical second stage model might consist of an equation for each coefficient. For example:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j + u_{1j}$$

In this example, the intercept (β_{0j}) is hypothesized to be a function of the overall mean of the outcome variable (γ_{00}), a group characteristic (W_j), and a unique (or random) effect associated with each group (u_{0j}). The slope (β_{1j}) is hypothesized to be a function of the mean of the slopes across groups, the effect of some group characteristic (γ_{11}), and a unique (or random) effect associated with each group. Here, the slope is considered to be *random*, since u_{0j} is included in the model. The slope would be considered to be *fixed* if this latter term were omitted from the model. Multilevel models can be expanded to a third stage that might examine the effects of an overarching unit (e.g., office building, school district) on the coefficients from the second stage model.

The various pieces of the second stage model can be substituted into the equation for individuals to yield a single equation that simultaneously explains between-group variance and within-group variance. An example follows:

$$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + \gamma_{10} (X_{ij}) + \gamma_{11} W_j (X_{ij}) + u_{0j} + u_{1j} (X_{ij}) + r_{ij}.$$

Here γ_{10} estimates the within group effect and γ_{01} represents the between group effect. γ_{11} is a cross-level interaction that measures the effect for a given person in a given group. γ_{00} is the intercept; the remaining terms and are residuals associated with individuals (r_{ij}) and with the parameters (u_{0j})

Centering in Multilevel Models

In multiple regression, the intercept is defined as the expected value of the dependent variable when the value of the predictor is 0 (Cohen and Cohen, 1983). A similar interpretation can be made with regard to multilevel models. That is, for the Level 1 model, the value of the intercept (β_{0j}) is defined as the expected value on the outcome measure (Y_{ij}) for an individual in group j with a value of 0 for X_{ij} (Bryk and Raudenbush, 1992). Since β_{0j} becomes the dependent variable in the Level 2 models, its meaning must be clear so the researcher can understand what is being predicted in the second stage. The key issue involves setting X_{ij} equal to 0, which in some situations, results in a nonsensical value for X.

For example, if X represents age, what would it mean to have an age of 0? Or, suppose that X is the score on a test that ranges from 10-100, what does it mean to have a test score of 0? Since it is often meaningless to have values of 0 for X, centering is used to scale the values of X and to purposefully place the value of 0 at a meaningful point. The literature on centering (Bryk and Raudenbush, 1992; Hofmann and Gavin, 1995; Kreft, de Leeuw, and Aiken, 1995; Schumacker and Bembry, 1997) focuses primarily on four methods for scaling the predictor variables: (1) the natural or raw score metric, also referred to as *no centering*; (2) grand mean centering; (3) group mean centering; and (4) centering on specific selected values.

Scaling on the Raw Score Metric

Under this method, variables are left in their original form. The meaning of β_{0j} in this case is the expected value for Y_{ij} when $X_{ij} = 0$. This metric may be meaningful in cases where 0 has a real value, such as when X measures hours of instruction or training, and 0 might indicate no instruction/training. Additionally, when X represents a dummy coded variable, then β_{0j} will represent the expected value for the individuals with dummy code values of 0. For instance, if individuals are coded so that African Americans have a value of 1 and Caucasians a value of 0, β_{0j} will represent the expected outcome value for Caucasians.

Grand Mean Centering

With grand mean centering, each X value is expressed as it deviation from the variable's grand mean, noted as $(X_{ij} - X_{..})$. This approach to centering anchors the meaning of X at the grand mean for the sample under study. With grand mean centering, B_{0j} is the expected outcome for subjects whose value on X_{ij} is equal to the grand mean. Bryk and Raudenbush (1992) point out that with grand mean centering, the intercept can be interpreted as an adjusted mean, in the same way one thinks of an adjusted mean with analysis of covariance (ANCOVA) models. In this case, the intercept would be thought of as:

$$\beta_{0j} = \overline{y}_j + \beta_{1j} (X_{ij} - X_{..}).$$

As with ANCOVA, grand mean centering allows consideration of an effect after partialling out or controlling for other effects. The variance, τ_{0j}, of β_{0j} is then the variance in the adjusted means.

Although grand mean centering is typically considered in connection with continuous variables, it can also be used for dummy coded variables. Grand mean centering of dummy coded variables allows the dummy coded variable to take on two values. The grand mean for the X variable will equal the proportion of individuals coded as 1. Consider the example where African Americans are coded as 1 and Caucasians as 0. The grand mean for X will be the proportion of African Americans in the sample. If a particular individual is African American, the grand-mean centered value will equal $X_{ij} - X_{..}$ (or 1 minus the proportion of African Americans), which is the proportion of Caucasians. For Caucasians, the grand mean centered variable will equal 0 minus the proportion of African Americans, resulting in a negative value, or minus the proportion of African Americans.

Group Mean Centering

Group mean centering expresses each X value as its deviation from the particular group's mean. That is, the value of X for an individual in group j would be her/his deviation from group j's mean on X, expressed as $(X_{ij} - X_{.j})$. Here, the value of β_{0j} is the unadjusted mean for group j, thus the

expected outcome for an individual depends on the group to which the individual belongs. The variance of the unadjusted means is simply the observed variance about the group means on *X*. One advantage of group mean centering is that it maintains orthogonality between level 1 and level 2 models (Bryk and Raudenbush, 1992), often making it easier to obtain a converged solution (Fein and Lissitz, 2000).

As with grand mean centering, group mean centering can be used for dummy coded variables. Continuing with the above example (African Americans coded 1, Caucasians coded 0), the group mean will be the proportion of African Americans in the group. For African Americans, the group-centered, dummy-coded value will be the proportion of Caucasians in group *j*; for Caucasians, *X* will be negative and will equal minus the proportion of African Americans in group *j*.

Centering on Specific Values

Sometimes there exists a specific value for *X* about which the researcher is interested. The value might be based on theory, a population mean, or possibly some baseline or cutoff limit. This type of centering operates much like grand mean centering, since it essentially involves adding or subtracting a constant from each case. With this method, each individual's score is expressed as a deviation from the specific value, not from the grand mean derived from the sample at hand. β_{0j} is then interpreted as the expected outcome for individuals who score at this preset *X* value.

Centering the Level 2 Predictors

Bryk and Raudenbush (1992) indicate that centering the level 2 predictors (the W's) is not as critical an issue as centering the level 1 predictors. Interpretation of the intercepts for the level 2 equations does not rely on the metric chosen for the level 2 predictors. Level 2 predictors may be centered to make them more easily interpretable. Or, they may be centered when an interaction variable will be used, to reduce the collinearity among variables.

Implications of Centering Choices

Studies of centering find that it enhances the interpretation of results and reduces the correlation between intercept and slope estimates across groups. This multicollinearity can cause convergence problems in obtaining a solution (Bryk and Raudenbush, 1992). Comparisons of centering methods indicate that raw metric scaling and grand mean centering produce nearly equivalent results (Burton, 1993; Kreft, de Leeuw, and Aiken, 1995). Since grand mean centering involves a linear transformation of the values for the centered variables, it will change the value of the intercept but not the slopes.

Group mean centering, on the other hand, appears to produce results that differ from those based on other centering options. This is primarily because group mean centering is not a simple linear transformation of the variables. Instead, it essentially introduces a new variable (Hofmann and Gavin, 1998). Whereas grand mean centering subtracts a constant value from the variable for all individuals, group mean centering subtracts a different value depending on the group the individual is in. Research on centering options has shown that group mean centering produces results that differ from other approaches and introduces the potential for misspecification if the group mean is not included as a predictor of the intercept at level 2 (Kreft, et al., 1995; Cohen, Rathbun, and Krotki, 1997).

Group mean centering has also been shown to alter the conclusions the researcher might draw about the relative importance of variables in the model (Burton, 1993; Hofmann et al., 1997). Variables that are not statistically significant in a group-mean centered model may be statistically significant in a grand mean centered model. Burton (1993) focused on this issue, studying the relationships between minority status and math achievement under varying centering options. His findings indicated that using raw metric scaling or grand mean centering of the minority status variable led to the conclusion that only the *individual student's* minority status had an effect on student's mathematics achievement. Use of group mean centering resulted in a statistically significant effect associated with the average *minority status of the school*. Although group mean centering can introduce complexity into model specification and interpretation, Bryk and Raudenbush (1992) and others (Kreft et al., 1995; Schumacker and Bembry, 1997; Hofmann and Gavin, 1997) find it to be exceedingly useful for studying contextual effects.

Numerical Examples

Description of Data Set and Variables

For the examples in this paper, the Level 1 dependent variable is a measure of student achievement in NELS:88 based on a test of math and reading (F12XCOMP) given to 10th graders (mean=51.55, sd=10.01, range=30.31 to 71.82). The independent variables used in the examples include ethnicity; the socioeconomic condition of the family; and perceived support from the teacher. The examples were based on 12,652 individuals and 657 schools, with an average of 20 students per school (mode=21; range=10 to 74).

Ethnicity is a categorical variable in NELS:88 that was dummy coded as follows: Asian Pacific Islanders and white non-Hispanic students were coded as 1 (n=2,515) and all other groups were coded as 0. Socioeconomic status (SES) is based on a composite variable available in the NELS:88 database (mean=.05, sd=.80, range=-3.29 to 2.76). Perceived teacher support is a factor constructed as part of an earlier study (information available from first author upon request) from a principal components analysis of questions asking for students' perceptions about the teachers in their school. Five questions were included in the factor:

- students get along well with teachers
- the teaching is good at this school
- teachers are interested in students
- when I work hard, teachers praise my efforts
- most teachers listen to me

Students responded to each question using a 5-point rating scale. In constructing the factor, individual questions were given approximately equal weight (alpha = .64). Values for the factor were standardized to have a mean of 0 and a standard deviation of 1.

School level explanatory variables included the school mean values for each of the Level 1 independent variables; that is, the average SES for students in a school; the proportion of non-minority (Asian and white) students in the school; and the school mean value for the Perceived Support factor. In addition, a factor constructed from principal components of questions asked of

the school administrators (as part of earlier study referenced above) was used to represent School Climate. Three questions were included in the factor:

- there is a positive relationship between school and parents.
- teachers press students to achieve
- students are expected to do homework

Administrators responded to each question using a 5-point rating scale. In constructing the factor, individual questions were given approximately equal weight (alpha=.65). Values for the factor were standardized to have a mean of 0 and a standard deviation of 1.

Demonstration 1: Basic Model

This demonstration is based on a model with three predictors included in the level 1 model (SES, ethnicity, and teacher support) and no predictors in the level 2 model. The slope for β_{2j} was treated as a fixed effect, while other parameters were treated as random. This decision was based on preliminary runs that indicated that the random variance component for β_{2j} was not statistically significant. The model appears below.

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}\ (SES) + \beta_{2j}\ (RACE) + \beta_{1j}\ (TCHRSUPP) + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + u_{3j}$

This results in the following combined model:

$$Y_{ij} = \gamma_{00} + \gamma_{10}\ (SESij) + \gamma_{20}\ (RACE_{ij}) + \gamma_{30}\ (TCHRSUPP_{ij}) + u_{0j} + u_{1j} + u_{3j} + r_{ij}.$$

Before comparing the results for different types of centering, it is instructive to consider the interpretation of the various coefficients:

γ_{00} is the intercept, or the value for achievement when all predictor variables equal zero.
γ_{10} is the average slope across schools when individuals' SES is used to predict achievement
γ_{20} is the average slope when individuals' Race is used to predict achievement
γ_{30} is the average slope when individuals' perception of Teacher Support is used to predict achievement
u_0 is the between-group variance associated with the intercepts (also referred to as tau_0)
u_1 is the between-group variance associated with the slopes for SES (tau_1)
u_3 is the between-group variance associated with the slopes for Teacher Support (tau_3)
r_{ij} is the unexplained variance associated with the level-1 model (within-group variance)

Table 1 displays the results from analyses using three centering options. The table includes estimates for the parameters, their standard errors (se), reliability estimates for the coefficients, and the variance components for the between-group and within-group variance.

Table 1. Comparisons of Three Centering Options for the Basic Model.

Parameter	Raw Metric	Grand Mean Centering	Group Mean Centering
γ_{00} (se)	47.93 (.21)*	51.54 (.12)	51.44 (.21)
γ_{10} (se) (SES)	4.67 (.11)	4.67 (.11)	4.12 (.12)
γ_{20} (se) (RACE)	4.18 (.22)	4.18 (.22)	3.85 (.25)
γ_{30} (se) (TCHSUPP)	1.66 (.08)	1.66 (.09)	1.53 (.09)
Reliability of B_0	.42	.43	.85
Reliability of B_1	.04	.04	.02
Reliability of B_3	.10	.10	.10
Var Comp u_0	4.98, p=.00, df=654	5.04, p=.00, df=654	23.86, p=.00, df=654
Var Comp u_1	.45, p=.018	.45, p=.018	.15 , p=.06
Var Comp u_3	.56, p=.042	.56, p=.042	.57 , p=.046
Var Comp r_{ij}	65.20	65.20	65.11

*All coefficients were statistically significant, p<.01.

Results are explained for the raw metric example. Here, the average achievement score for the sample was 47.93 (with a standard error of .21). Each unit increase in SES is associated with an increase in achievement of 4.67 points; likewise each unit increase in teacher support results in an increase of 1.66 points in achievement. Because race is a dummy coded variable, the coefficient for race represents the difference in performance between minority (coded 0) and non-minority students (coded 1). On average, non-minority students scored 4.18 points higher than minority students.

Comparison of the coefficients for the three types of centering methods shows that results are very similar for raw metric and grand mean centering. Raw metric and grand mean centering differed only in estimates of the intercepts. Group mean centering produced different intercept and slope parameters. The variance component for the intercepts, u_0, was substantially larger with group mean centering than for the other approaches.[1]

Reliability estimates are provided for the tau's for each random effect. In HLM, reliability indicates that percentage of *tau* that is reliable parameter variance. The total variance consists of both parameter variance and sampling variance. Reliability thus estimates how much of the total variance can be explained by the between-group model(s) (Arnold, 1992). The formula for estimating reliability is:

(parameter variance) / (parameter variance + error variance).

HLM estimates the reliability of a given parameter for each group-level sample, where error variance depends on the within-group sample size. The overall measure of reliability is the average of the within-group reliabilities (Bryk and Raudenbush, 1992). The reliability estimate of the intercept was distinctly higher for group mean centering than for other approaches but slightly less for the SES slope.

[1] To examine whether this finding was peculiar to this particular data set, we ran these analyses on two data sets routinely available with the HLM software (the *High School and Beyond* and the *Vocabulary* data sets.) For both cases, group mean centering produced larger estimates of tau for the intercept than did other centering approaches.

Estimating Explained Variance

Because hierarchical linear modeling appears to closely parallel linear multiple regression, it seems reasonable to expect that a statistic like R^2 (percent of variance explained) should be available. This statistic is not routinely provided with HLM software, but formulas have been developed to estimate it. These formulas estimate explained variance by comparing reductions in error variance for series of nested models (Arnold, 1992; Kreft and De Leeuw, 1998; Snijders and Bosker, 1999). A complication arises, however, because such calculations depend on having an estimate of total variance (e.g., within-group variance + between-groups variance). Kreft and De Leeuw, 1998 (p. 116) explain that with HLM the within-group variance (r_{ij}) and the between-group variance (tau_0) do not sum to total variance due to confounding (the level-1 coefficients cannot be separated into between and within parts). As a result, sometimes adding variables to the model can increase between-groups variance and decrease the amount of explained variance – a counterintuitive finding. Snijders and Bosker (1999) suggested that such a finding could indicate that the model is misspecified. For instance, decreases in variance explained can occur when basic assumptions are violated (e.g., level 1 or level 2 errors are correlated with one or more X variables), which can happen when important variables are not included in the model. They noted that R^2 can be helpful as a diagnostic tool to signify misspecified models.

For each of the demonstrations in this paper, we provide estimates of R^2 to illustrate how the statistic might be calculated. We wish to point out, however, that reporting R^2 for multilevel models is controversial. While Kreft and de Leeuw (1998:119) provided formulas, they also concluded their discussion by saying that the concept of R^2 in multilevel models is "ill defined and ambiguous," and the usefulness of the statistic is limited to random intercept models. Some authors (such as Goldstein, personal communication; and recent discussions on a multilevel modeling listserve) discourage its use altogether.

Kreft and de Leeuw (1998) and Snijders and Bosker (1999) proposed different formulas for calculating explained variance. The Kreft and de Leeuw formula is:

For level 1:

(1) $R^2_{1KD} = (\sigma^2 \text{ original model} - \sigma^2 \text{ new model}) / \sigma^2 \text{ for original model}$

For level 2:

(2) $R^2_{2KD} = (\tau^2 \text{ original model} - \tau^2 \text{ new model}) / \tau^2 \text{ for original model}$

where σ^2 is within group variance (r_{ij}) and τ^2 is between group variance (u_{0j}). The Snijders and Bosker formula is:

For level 1:

(3) $R^2_{1SB} = 1 - [(\sigma^2 \text{ new model} + \tau^2 \text{ new model}) / (\sigma^2 \text{ for original model} + \tau^2 \text{ original model})]$

For level 2:

(4) $R^2_{2SB} = 1 - [(\sigma^2 \text{ new model}/n + \tau^2 \text{ new model}) / (\sigma^2 \text{ for original model}/n + \tau^2 \text{ original model})]$

The n in equation 4 is intended to be a measure of the size of the group (e.g., for each class or school). In balanced designs, the group size is consistent across groups and this value can be used for n. In unbalanced designs, deciding on the appropriate value of n is not as straightforward. Snijders and Bosker (1999) suggest using either a measure of the average group size for the sample or an estimate of the typical group size in the population. They point out that when the group size in the population is very large, the value for the within groups variance will be diminished, and R^2 will simply be a ratio of the two estimates of between group variance.

Both pairs of authors indicate that these formulas are for models with random intercepts and do not apply for models with random slopes. However, Snijders and Bosker (1999) proposed that estimates of R^2 for models with random slopes can be obtained by re-running the models with fixed slopes and using the values for between and within groups variance to estimate R^2's for the random slopes model. This was done for the models in the present study, and R^2 estimates were calculated using both formulas. The fully unconditional model was run to obtain the between group (τ^2=23.04) and within group (σ^2=77.59) variance components. Results appear below. For these calculations, the modal value for group size (21) was used as an estimate of n.

Table 2. Comparison of Incremental R^2 for Different Centering Methods for the Basic Model

Type of Centering	Within Groups σ^2	Between Groups τ^2	Level 1		Level 2	
			R^2_{1KD}	R^2_{1SB}	R^2_{2KD}	R^2_{2SB}
Raw metric	65.85	5.17	.1513	.29	.78	.69
Grand mean	65.85	5.17	.1513	.29	.78	.69
Group mean	65.68	23.82	.1534	.11	-.03	0

Several observations can be made about these results. First, as might be expected, the R^2 values for raw metric and grand mean centering are identical in all cases. Second, with the exception of the level 1 estimate for the KD formula, the explained variance is higher for raw metric and grand mean centering than for group mean centering. It is also noteworthy that the two formulas produced different estimates.

Demonstration 2: Intercepts-as-Outcomes Model

For this demonstration, an intercepts-as-outcomes model was run. The same three predictor variables were included in the level-1 model (SES, ethnicity, and teacher support), and their mean values were included in the Level 2 model as predictors of the intercept. The β_{2j} slope was treated as a fixed effect. The equations for this model appear below:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}\ (SES_{ij}) + \beta_{2j}\ (RACE_{ij}) + \beta_{1j}\ (TCHRSUPP_{ij}) + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}\ (MEANSES) + \gamma_{02}\ (MEANTSUPP) + \gamma_{03}\ (MEANRACE) + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30} + u_{3j}$

Combining the Level 1 and 2 models results in the following:

$$Y_{ij} = \gamma_{00} + \gamma_{01}\,(\text{MEANSES}) + \gamma_{02}\,(\text{MEANTSUPP}) + \gamma_{03}\,(\text{MEANRACE}) + \gamma_{10}\,(\text{SES}_{ij}) + \gamma_{20}\,(\text{RACE}_{ij}) + \gamma_{30}\,(\text{TCHRSUPP}_{ij}) + u_{0j} + u_{1j} + u_{3j} + r_{ij}.$$

The difference between this model and the model in the first example is the addition of predictor variables for the intercept. It is useful to consider the interpretation of the additional coefficients before comparing the effects of centering options. As shown by the level 2 model:

γ_{00} is the average achievement level across schools; the average value for β_{0j}.
γ_{01} is the change in the intercept β_{0j} associated with mean SES for the school.
γ_{02} is the change in the intercept β_{0j} associated with an increase in mean Teacher Support.
γ_{03} is the change in the intercept β_{0j} associated with the proportion of non-minority students in the school.

When the level 2 terms are substituted into the level 1 equation, it is possible to examine the effects of individual characteristics as compared to group characteristics on individuals' achievement. For example, γ_{01} represents the group effect of SES while γ_{10} represents the individual effect.

Table 3 displays the results from analyses using three centering options and provides similar information as that included in Table 1.

Table 3. Comparison of Centering Options for Means-As-Outcomes Model

Parameter	**Raw Metric**	**Grand Mean Centering**	**Group Mean Centering**
γ_{00} (se)	47.72 (.38) *	51.04 (.43)	47.71 (.38)
γ_{01} (se) (Mean SES)	2.46 (.28)	2.46 (.28)	6.56 (.25)
γ_{02} (se) (Mean TSUPP)	1.07 (.31)	1.07 (.31)	2.61 (.30)
γ_{03} (se) (Mean RACE)	.52 (.53), NS	.53 (.53), NS	4.38 (.46) Sig
γ_{10} (se) (SES)	4.13 (.13)	4.13 (.13)	4.13 (.12)
γ_{20} (se (RACE)	3.85 (.25)	3.85 (.25)	3.85 (.25)
γ_{30} (se) (TSUPP)	1.53 (.09)	1.53 (.09)	1.53 (.09)
Reliability of B_0	.36	.36	.50
Reliability of B_1	.02	.02	.01
Reliability of B_3	.11	.11	.11
Var Comp u_0	3.75, df=651,p=.00	3.73, p=.00	3.85, p=.00
Var Comp u_1	.20, df=654, p=.059	.20, p=.059	.09, p=.062
Var Comp u_3	.63, df=654, p=.045	.63, p=.045	.59, p=.047
Var Comp r_{ij}	65.07	65.07	65.15

* All coefficients were significant, p<.01 except where noted as "NS."

Comparison of the coefficients obtained for the three centering options shows that, again, results are nearly identical for raw metric scaling and grand mean centering, with the exception of the intercepts (as expected). The coefficients for the predictors of the intercepts differ for group mean centering versus the other two options. The differences are quite large and influence the interpretations of the findings.

The values for Mean SES and Mean Teacher Support are higher for the group mean centered model than for the others. Note, for example, that with grand mean centering, the coefficient for Mean SES of the school is 2.46, while the coefficient for individuals is 4.13, roughly twice that for schools. For the group centered model, the reverse is true; the coefficients are 6.56 for schools and 4.13 for individuals. With grand mean centering, the results indicate that the SES for individuals is more important than the average SES of the school for predicting individual's achievement. With group mean centering, the results suggest that the school setting is more important.

While the values for the SES coefficients changed under the different centering options, they all remained statistically significant. This was not the case for the ethnicity variable, however. Under grand mean centering and the raw metric approach, the coefficient for average proportion of non-minorities in a school was approximately .52, which is not statistically significant. For group mean centering, the coefficient was more than eight times higher at 4.38, which is statistically significant. Under grand mean centering, the results suggest that there were differences in achievement for the two ethnic groups but that the ethnic composition of the school did not make a difference. Under group mean centering, the results indicate a statistically significant effect for the school's ethnic composition.

Estimates of Explained Variance

Estimates of variance explained by this model appear in Table 4. Here, the estimate was calculated to represent the increment in variance explained by the intercepts-as-outcomes model as compared to the basic model (presented in demonstration 1). Thus, the original values of σ^2 and τ^2 used in the calculations are those that appear in Table 2.

Table 4. Comparison of Incremental R^2 for Different Centering Methods for Intercepts-as-Outcomes Model

Type of Centering	Within Groups σ^2	Between Groups τ^2	Level 1		Level 2	
			R^2_{1KD}	R^2_{1SB}	R^2_{2KD}	R^2_{2SB}
Raw metric	65.69	3.82	.002	.02	.26	.16
Grand mean	65.69	3.82	.002	.02	.26	.16
Group mean	65.69	3.82	-.0002	.22	.83	.74

Again, results were identical for raw metric and group mean centering; and the results differed for the two methods for calculating R^2. An instance of a negative increase was noted for the group mean centered model when the Kreft and de Leeuw formula was used. This occurred because the within-groups variance increased from 65.68 for the basic model to 65.69 for the intercepts-as-outcomes model. While the increase in within groups variance is small (as is the associated decrease in explained variance), it may be attributable to model misspecification, given

that the random components of the slopes, u_1 and u_3, which were statistically significant, were omitted from the models in order to calculate R^2. For all but the R^2_{1KD} estimates, group mean centering resulted in a larger increase in explained variance.

Demonstration 3: Intercepts- and Slopes-as-Outcomes Model

For the next demonstration, an intercepts and slopes as outcomes model was run. The same Level 1 model was run as in the previous demonstration, but Mean SES was included in the Level 2 model as a predictor of the SES slope.

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}\ (SES_{ij}) + \beta_{2j}\ (RACE_{ij}) + \beta_{1j}\ (TCHRSUPP_{ij}) + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}\ (MEANSES) + \gamma_{02}\ (MEANTSUPP) + \gamma_{03}\ (MEANRACE) + u_{0j}$
$\beta_{1j} = \gamma_{10} + \gamma_{11}\ (MEANSES) + u_{1i}$
$\beta_{2j} = \gamma_{20}$
$\beta_{3j} = \gamma_{30} + u_{3j}$

Combining the Level 1 and 2 models results in the following:

$$Y_{ij} = \gamma_{00} + \gamma_{01}\ (MEANSES) + \gamma_{02}\ (MEANTSUPP) + \gamma_{03}\ (MEANRACE) + [\gamma_{10} + \gamma_{11}\ (MEANSES) + u_{1i}](SES_{ij}) + \gamma_{20}\ (RACE_{ij}) + \gamma_{30}\ (TCHRSUPP_{ij}) + u_{0j} + u_{3i} + r_{ij}.$$

Combining terms yields the following equation:

$$Y_{ij} = \gamma_{00} + \gamma_{01}\ (MEANSES) + \gamma_{02}\ (MEANTSUPP) + \gamma_{03}\ (MEANRACE) + \gamma_{10}\ (SES_{ij}) + \gamma_{11}\ (MEANSES)\ (SES_{ij}) + \gamma_{20}\ (RACE_{ij}) + \gamma_{30}\ (TCHRSUPP_{ij}) + u_{0j} + u_{1i}\ (SES_{ij}) + u_{3i} + r_{ij}.$$

The difference between this model and the model in the second example is the addition of a predictor variable for the MEANSES slope. The addition of an explanatory variable for the slope allows the researcher to consider if the relationship between the X variable (individual SES) and the outcome (achievement) varies depending on the average SES for the school. In the level 2 model, the additional coefficients would be interpreted as follows:

γ_{10} is the average slope when individuals' SES is used to predict achievement

γ_{11} is the change in β_{1j} associated with the average SES for the school.

As described above, when the level 2 terms are substituted into the level 1 equation, it is possible to examine the effects of individual versus group characteristics on individuals' achievement and to see the cross-level interaction effects. In the combined equation, γ_{11} represents the cross-level interaction and can be interpreted in the way interactions are generally interpreted. In this example, the cross-level interaction estimates the effect on achievement of particular combinations of school SES and individual SES, beyond the effects of school SES alone and individual SES alone. Table 5 displays the results from analyses using three centering options.

Table 5. Comparison of Centering Options for Intercepts- and Slopes-As-Outcomes Model*

Parameter	Raw Metric	Grand Mean Centering	Group Mean Centering
γ_{00} (se)	47.44 (.39)	50.79 (.44)	47.71 (.38)
γ_{01} (se) (Mean SES)	2.26 (.29)	2.28 (.29)	6.56 (.25)
γ_{02} (se) (Mean TSUPP)	.93 (.32)	.93 (.32)	2.61 (.30)
γ_{03}(se) (Mean RACE)	.69 (.53)NS	.69 (.53) NS	4.38 (.46) Sig
γ_{10} (se) (SES)	4.12 (.12)	4.12 (.13)	4.13 (.12)
γ_{11} (se) (SES) (Mean SES)	.51 (.21) Sig	.51 (.21) Sig	.23 (.26) NS
γ_{20} (se) (RACE)	3.87 (.25)	3.87 (.25)	3.86 (.25)
γ_{30} (se) (TSUPP)	1.53 (.09)	1.53 (.09)	1.53 (.09)
Reliability of B_0	.36	.36	.50
Reliability of B_1	.03	.03	.01
Reliability of B_3	.11	.11	.10
Var Comp u_0	3.69, df=651,p=.000	3.67, p=.000	3.85, p=.000
Var Comp u_1	.27, df=653, p=.054	.27 , p=.054	.12, p=.060
Var Comp u_3	.64, df=654, p=.044	.63, p=.044	.57, p=.047
Var Comp r_{ij}	65.04	65.05	65.14

* All coefficients were significant, p<.01 except where noted as "NS."

This model added a predictor variable for the SES slope, γ_{11}, which estimates the cross-level interaction effect. This coefficient is the term that was affected by the centering method. For raw metric scaling and grand mean centering, the coefficient for γ_{11} indicates that a unit increase in the average SES combined with a unit increase in individual SES produces a .51 increase in achievement. With group mean centering, the coefficient is half as large (.23). With raw metric scaling or grand mean centering, the cross-level interaction is statistically significant; but with grand mean centering, the interaction is not significant.

Estimates of Variance Explained

Estimates of variance explained by this model appear in Table 6. Again, the estimates were calculated to represent the increment in variance explained by the intercepts- and slopes-as-outcomes model as compared to the intercepts-as-outcomes model (presented in demonstration 2). Thus, the original values of σ^2 and τ^2 used in the calculations were those that appear in Table 4.

Table 6. Comparison of Incremental R^2 for Different Centering Methods for Intercepts-and Slopes-as-Outcomes Model

Type of Centering	Within Groups σ^2	Between Groups τ^2	Level 1		Level 2	
			R^2_{1KD}	R^2_{1SB}	R^2_{2KD}	R^2_{2SB}
Raw metric	65.69	3.77	.000	.0007	.013	.007
Grand mean	65.69	3.77	.000	.0007	.013	.007
Group mean	65.69	3.83	.000	-.0001	-.003	-.001

These results show that increases in explained variance were very small for all three centering methods, and there were two instances where the increase was negative. Again the values for R^2 were different for group mean centering versus the other two approaches.

Using Centering to Address Specific Research Questions

The second part of the paper explores ways to use centering to address specific research questions. The first section focuses on ways to examine contextual effects, and the second section deals with cross-level interaction effects.

Studying Contextual Effects

Bryk and Raudenbush (1992) point out that researchers can use HLM to evaluate contextual effects, that is, when the aggregate of a person-level characteristic is related to the outcome after controlling for individual characteristics. Such models require that the aggregate value be included as a predictor of the intercept (Cohen et al., 1997).

Grand mean and group mean centering were compared using the dataset described above and the following intercepts-as-outcomes model:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij}) + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01} X_{.j} + u_{0j}$

$\beta_{1j} = \gamma_{10}$

Under grand mean centering, the combined model would be:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(X_{.j}) + \gamma_{10}(X_{ij} - X_{..}) + u_{0j} + r_{ij}$$

Here, γ_{01} represents the contextual effect and γ_{10} is the individual effect. Under group mean centering, the combined model would be:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(X_{.j}) + \gamma_{10}(X_{ij} - X_{.j}) + u_{0j} + r_{ij}$$

Determining the contextual effect requires combining like terms and subtracting:

$$Y_{ij} = \gamma_{00} + (\gamma_{01} - \gamma_{10})(X_{.j}) + \gamma_{10}(X_{ij}) + u_{0j} + r_{ij}$$

Now, the contextual effect is represented by $\gamma_{01} - \gamma_{10.}$ If the researcher were to use group mean centering and omit the aggregate value as a level-2 predictor, the following combined equation would be obtained:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij}) - \gamma_{10}(X_{.j}) + r_{ij} + u_{ij}$$

This model suggests that the effect of the group value for X is equal and *exactly opposite* of the effect of individual's value on X, a finding that does not make sense.

Numerical Example—Demonstration 4

To demonstrate the counterintuitive result that can occur from this type of misspecification, the teacher support variable was used to predict scores on the reading and math achievement test by running two different models. For Model 1, the school mean was not included as a level 2 predictor; for Model 2 it was. Teacher support was group mean centered for both models. Thus, Model 1 was:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - X_{.j}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{ij}$$
$$\beta_{1j} = \gamma_{10}$$

where X_{ij} is teacher support. By substitution:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij}) - \gamma_{10}(X_{.j}) + r_{ij} + u_{ij}$$

Model 2 was:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - X_{.j}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(X_{.j}) + u_{ij}$$
$$\beta_{1j} = \gamma_{10}$$

And, by substitution:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(X_{.j}) + \gamma_{10}(X_{ij} - X_{.j}) + r_{ij}$$

For Model 1, the intercept was equal to 51.45 and γ_{10} was equal to 1.54, resulting in the following equation.

$$Y_{ij} = 51.45 + 1.54\,(X_{ij}) - 1.54\,(X_{.j})$$

The interpretation of these coefficients is that achievement increases by 1.54 for each unit increase in an individual's value of perceived teacher support. But, for each unit increase in the school mean value for teacher support, achievement scores *decrease* by 1.54. Thus, a student who felt supported in an unsupportive environment would be expected to do better than the student who felt supported in a supportive environment, a rather odd finding.

The intercept and coefficient for γ_{10} were the same in Model 2 as in Model 1. But, in Model 2, a value of 5.62 was found for γ_{01}, resulting in the following prediction equation:

$$Y_{ij} = 51.45 + 5.62\,(X_{.j}) + 1.54(X_{ij} - X_{.j})$$

The interpretation here is that a unit increase in the school's mean level of teacher support is associated with a 5.62 increase in individual's achievement. And, for each unit increase above the school mean, individuals increased their achievement scores by 1.54 points.

The equations for Model 1 and 2 were then used to predict individuals' achievement scores, and the results were graphed. Factor scores on the Teacher Support factor were used to group individuals into quartiles. Predicted achievement scores for individual in the lowest and highest quartiles on the Teacher Support factor were included in the figures. Figure 1 shows the result for Model 1, and Figure 2 shows the results for Model 2. Both figures show that achievement scores are predicted to be higher for individuals who felt most strongly supported by their teachers than for those who felt the least level of teacher support. However, in Figure 1, the school mean level of teacher support is negatively related to achievement. In Figure 2, the relationship is positive.

[Insert Figures 1 and 2 about here]

This was replicated (just for verification purposes) by running two models in which achievement was predicted by SES. The model which omitted the mean from level 2 produced the following results:

$$Y_{ij} = 51.45 + 4.53(X_{ij}) - 4.53\,(X_{.j})$$

And, with the mean at level 2,

$$Y_{ij} = 51.09 + 8.26\,(X_{.j}) + 4.53\,(X_{ij} - X_{.j}).$$

Cross-Level Interactions—Demonstration 5

Other research may examine cross-level effects, that is interactions between level 1 and level 2 variables. Cross-level interactions indicate that the relationship between the outcome measure and a given level-1 variable differs over the values of a group-level variable. Cross-level interactions are modeled by incorporating variables in the level-2 model as predictors for level-1 slopes. For instance, in demonstration 3 above, Mean SES was included as a level-2 predictor for β_{1j}, the equation that modeled the relationship between individual's SES and achievement. Thus, the cross-level interaction would test if the relationship between an individual's SES and his or her achievement varied depending on the mean SES of the school the individual attended.

Hofmann and Gavin (1998) pointed out that often a model such as that used with demonstration 3 is proposed as a way for examining cross-level interactions. They note, however, that when such a model is used along with grand mean centering, it is impossible to disentangle between-group and within-group effects for cross-level interaction. For example, the following model might be proposed:

Level 1:	$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij}) + r_{ij}$
Level 2:	$\beta_{0j} = \gamma_{00} + \gamma_{01} W_{.j} + u_{0j}$
	$\beta_{1j} = \gamma_{10} + \gamma_{11} W_{.j} + u_{0j}$

where *W* is a school level predictor variable. Under grand mean centering, the combined model would be (error terms omitted for simplicity):

$$Y_{ij} = \gamma_{00} + \gamma_{01}(W_{.j}) + \gamma_{10}(X_{ij} - X_{..}) + \gamma_{11} W_{.j}(X_{ij} - X_{..})$$

The cross-level interaction is represented by the final term, $W_{.j}(X_{ij} - X_{..})$. The problem is that the between-group and within-group effects cannot be partialled out; γ_{11} is a mix of the between-group and the within-group effects. Hofmann et al. (1998) show that this problem can be overcome by using the following model:

Level 1:	$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij}) + r_{ij}$
Level 2:	$\beta_{0j} = \gamma_{00} + \gamma_{01} X_{.j} + \gamma_{02} W_{.j} + \gamma_{03}(X_j W_j) + u_{0j}$
	$\beta_{1j} = \gamma_{10} + \gamma_{11} W_{.j} + u_{0j}$

Here, two predictor variables have been added to the level-2 model for the intercept. $X_{.j}$ is the group mean for variable X, which is needed under group mean centering as explained in demonstration 4. $X_j W_j$ is the term that makes it possible to disentangle the between- and within-groups effects. Under group mean centering, the combined model is (error terms omitted for simplicity):

$$Y_{ij} = \gamma_{00} + \gamma_{01} X_{.j} + \gamma_{02}(W_{.j}) + \gamma_{03}(XjW_{.j}) + \gamma_{10}(X_{ij} - X_{.j}) + \gamma_{11} W_{.j}(X_{ij} - X_{.j})$$

The relevant portions of the model are the terms $\gamma_{03}(X_j W_{.j})$ and $\gamma_{11} W_{.j}(X_{ij} - X_{.j})$. When multiplied through, these terms become:

$$\gamma_{03}(X_j W_{.j}) + \gamma_{11}(W_{.j} X_{ij}) - \gamma_{11}(W_{.j} X_{.j}).$$

Combining like terms yields:

$$(\gamma_{03} - \gamma_{11})(X_j W_{.j}) + \gamma_{11}(W_{.j} X_{ij}).$$

Here, $(\gamma_{03} - \gamma_{11})$ estimates the effect of the between-group interaction, while γ_{11} represents the within-group interaction.

Examination of the same model under grand mean centering shows that the researcher still cannot partial out the between-group and within-group effects of the interaction. Under grand mean centering the model is:

Level 1:	$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - X_{..}) + r_{ij}$
Level 2:	$\beta_{0j} = \gamma_{00} + \gamma_{01} X_{.j} + \gamma_{02} W_{.j} + \gamma_{03}(X_j W_{.j}) + u_{0j}$
	$\beta_{1j} = \gamma_{10} + \gamma_{11} W_{.j} + u_{0j}$

When combined:

$$Y_{ij} = \gamma_{00} + \gamma_{01} X_{.j} + \gamma_{02}(W_{.j}) + \gamma_{03} (XjW_{.j}) + \gamma_{10} (X_{ij} - X_{..}) + \gamma_{11} W_{.j} (X_{ij} - X_{..})$$

The relevant portions of the model are the terms $\gamma_{03} (XjW_{.j})$ and $\gamma_{11} W_{.j} (X_{ij} - X_{..})$. When multiplied through, these terms become:

$$\gamma_{03} (X_j W_{.j}) + \gamma_{11} (W_{.j} X_{ij}) - \gamma_{11} (W_{.j} X_{..}).$$

Here, there are no like terms to combine and the reader can see that γ_{11} is a mix of both between-group and within-group effects.

To demonstrate, a model was run using group mean centering in which W_j was the School Climate factor, X_j was the school level of SES, and $W_j X_j$ was the interaction of the two. The combined equation is repeated here for clarity:

$$Y_{ij} = \gamma_{00} + \gamma_{01} X_{.j} + \gamma_{02}(W_{.j}) + \gamma_{03} (X_j W_{.j}) + \gamma_{10} (X_{ij} - X_{.j}) + \gamma_{11} W_{.j} (X_{ij} - X_{.j})$$

Obtained values for the relevant coefficients were $\gamma_{01} = .41$, $\gamma_{02} = 7.79$, $\gamma_{03} = .36$, $\gamma_{10} = 4.11$, and $\gamma_{11} = -.05$. These values indicate: (1) a unit increase in the value for the School Climate is associated with a .41 increase in individuals' achievement; (2) a unit increase in the school's Mean SES is associated with a 7.79 increase in individuals' achievement; (3) each unit increase in an individual's SES in relation to the school's Mean SES is associated with a 4.11 increase in achievement; (4) a .05 decrease [2] in achievement is associated with a unit increase in individual SES combined with a unit increase in School Climate; and (5) a .41 (that is, .36 – -.05) increase in achievement is associated with the combination of a unit increase in School Climate and a unit increase in Mean SES. These latter two effects are the cross-level interactions.

Summary

In this paper, we presented five numerical examples to demonstrate the effects of centering choices on model parameters, explained variance, and interpretation of results. Demonstrations 1 through 3 show that results for raw metric scaling and grand mean centering tend to be similar and tend to differ markedly from results obtained under group mean centering.

In demonstration 1, the between-groups variance estimate was more than five times as large with group mean centering as with other methods. Hence, the increase in explained variance for level 2 (compared to the fully unconditional model) was smaller for group mean centering than for the other methods. In demonstration 2, parameter estimates varied considerably for raw metric/grand mean centering versus group mean centering; and one of the school-level variables (MEANRACE; representing proportion of whites and Asians in the school) was statistically significant for group mean centering but not for the other methods. Also in demonstration 2, the increase in R^2 for level 2 under group mean centering was substantially larger than for the other methods. In demonstration 3, the coefficient associated with the cross-level interaction was

[2] This coefficient was not statistically significant, indicating that there was essentially no effect for School Climate. We include it for demonstration purposes only.

statistically significant under raw metric/grand mean centering but not significant under group mean centering.

These types of differences lead to different interpretations of results, but these differences are only apparent when multiple centering methods are used and compared. If the researcher had chosen to use only one of the centering methods (as is more typically done), the types of interpretations would depend on the choice of centering method.

Centering is useful for a number of reasons. As shown in Demonstrations 4 and 5, it can be used to help disentangle and study between-group and within-group effects. Centering can also enhance the interpretation of results and reduce collinearity, but it can also alter the results and their interpretations. The advice offered by Kreft et al. (1995) is probably the most salient, "there is no statistically correct choice among centering options, but rather the choice should be driven by theory and by the intent of the research." The critical issue is that the researcher be aware of how centering decisions affect the interpretation of the results to avoid unknowingly drawing erroneous conclusions.

References

Arnold, C. (1992). An introduction to hierarchical linear models. *Measurement and Evaluation in Counseling and Development*, v.25, 58-90.

Burton, B. (1993). Some observations on the effect of centering on the results obtained from hierarchical linear modeling. Paper presented at the annual meeting of the American Educational Research Association, 1993, Atlanta, GA.

Bryk, A.S. and Raudenbush, S.W. (1992) *Hierarchical Linear Models*. Newbury Park, CA: SAGE Publications Inc.

Cohen, J and Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*.Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Rathbun, A., and Krotki, K. (1997). Hierarchical linear models and their application in educational research. Draft.

Fein, M.D. and Lissitz, R.W. (2000). Comparison of HLM and MLwiN multilevel analysis software packages: A Monte-Carlo investigation into the quality of the estimates. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Kreft, I., de Leeuw, J. (1998). *Introducing Multilevel Modeling*. Thousand Oaks, CA: SAGE Publications Inc.

Kreft, I., de Leeuw, J., and Aiken, L.S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research* 30, (1) 1-21.

Hofmann, D.A. and Gavin, M.D. (1998). Centering decisions in hierarchical linear models: implications for research in organizations. *Journal of Management*, 24 (5), 623-641.

Schumacker, R.E. and Bembry, K. (1997). Empirical characteristics of centering methods for level-1 predictor variables in HLM. *MLRV*, 23, 1-8.

Snijders, T.A.B. and Bosker, R.J. (1999). *Multilevel Analysis*. Thousand Oaks, CA: SAGE Publications Inc.

Figure 1. Group Mean Not Included

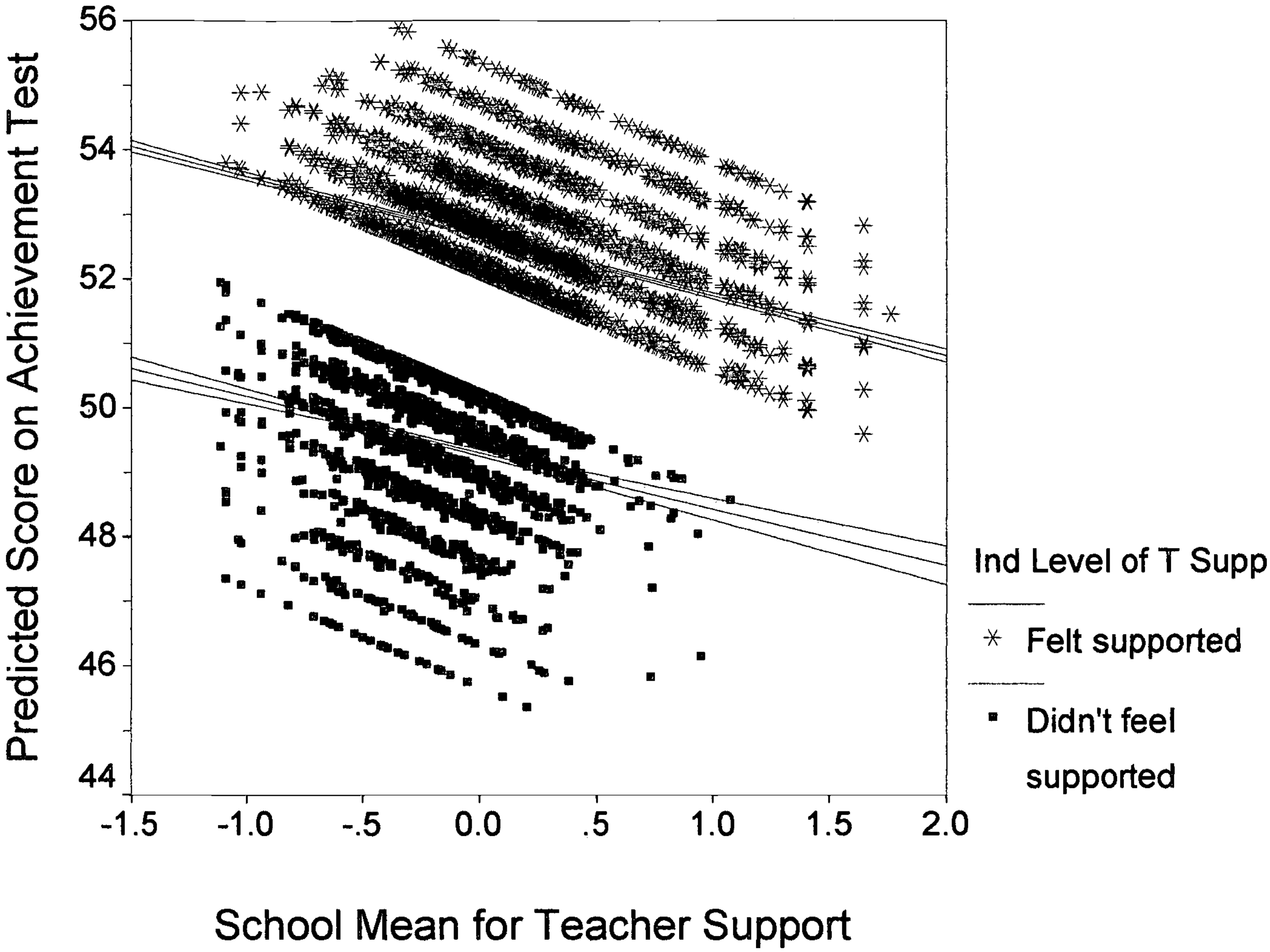
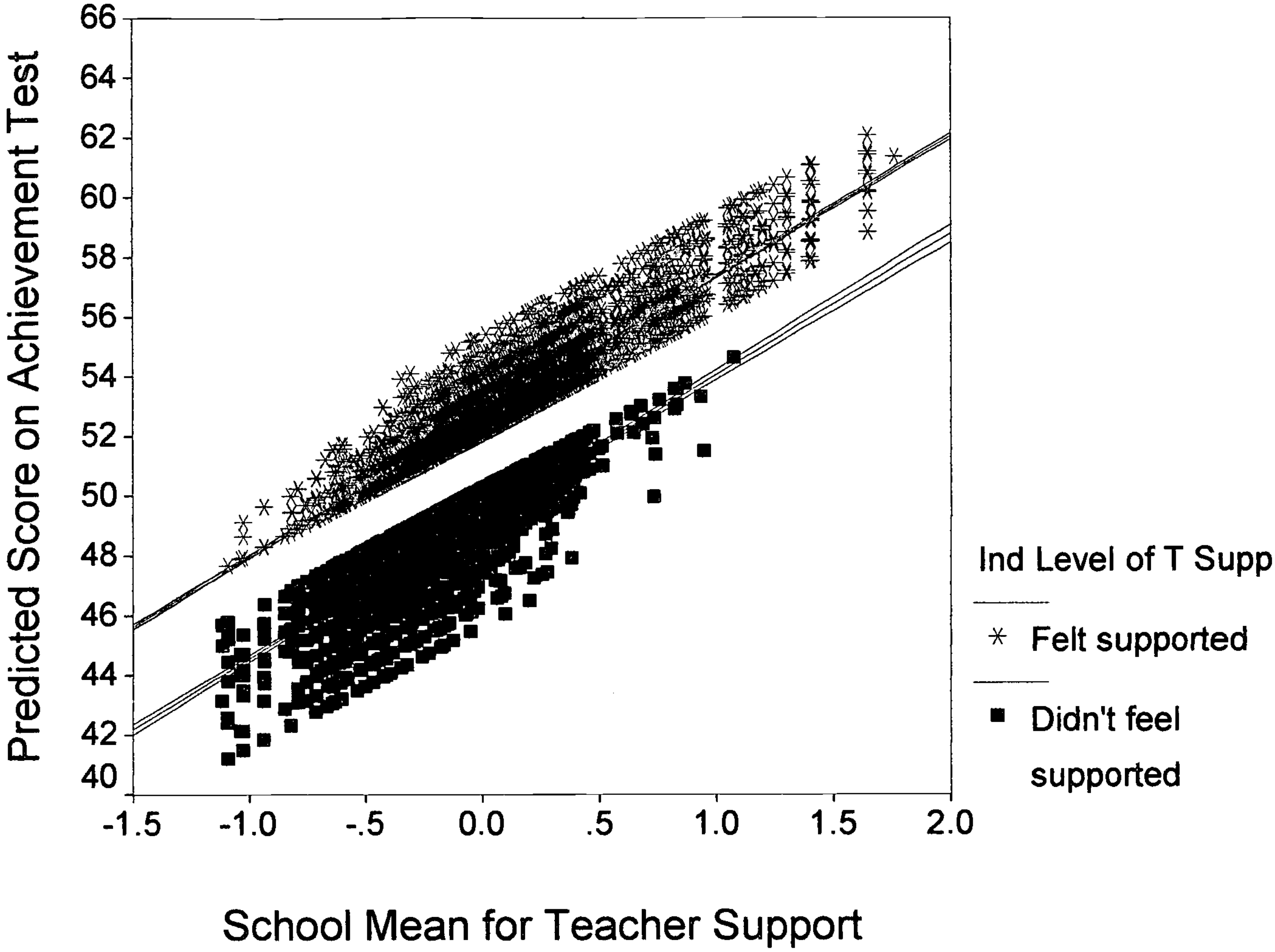


Figure 2. Group Mean Included



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

TM032798

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: The Effects of Centering on Parameter Estimates and Variance Explained in Multilevel Models

Author(s): Judith A. Koenig & Robert W. Lissitz

Corporate Source: University of Maryland

Publication Date: 4/2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☒	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Judith A Koenig

Printed Name/Position/Title: Judith A. Koenig, Program Officer

Organization/Address: National Research Council 2101 Constitution Ave, NW HA-450 Washington DC 20418

Telephone: 202-334-3399

FAX: 202-334-1453

E-Mail Address: JKOENIG@NAS.EDU

Date: 4/24/01

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)