15

Centering in Two-Level Nested Designs

James Algina
Department of Educational Phychology, School of Human Development and Organizational Studies in Education,
University of Florida, Gainesville, Florida

Hariharan Swaminathan
Department of Educational Psychology,
University of Connecticut, Storrs, Connecticut

15.1 CENTERING IN TWO-LEVEL NESTED DESIGNS

In this chapter we review basic results about centering independent variables in two-level nested designs. We present the major results in the context of an example in which the available data are math achievement (*MA*), socioeconomic status (*SES*), and the school a student attends. The goal of the data analysis is to demonstrate the effect of centering of the independent variable *SES* on the relationship between *SES* and the dependent variable *MA*. Three types of centering of *SES* are considered:

- Grand mean centering in which the grand mean of *SES* is subtracted from each *SES* score: $SES-\overline{\overline{SES}}$, where $\overline{\overline{SES}}$ is the average of all of the *SES* scores in the sample.
- Group mean centering in which for each student in a school, the school mean of *SES* is subtracted from the *SES* score: $SES-\overline{SES}_j$, where $\overline{SES}_j$ is the average of *SES* score for students who are in the analysis and attended school *j*.
- No centering, in which a mean is not subtracted from *SES*.

In our review, we emphasize that centering, in particular group mean centering, addresses certain potential assumption violations and also affects the interpretation of the fixed effects and variance components in the model. Because of the impact of centering decisions on interpretation, we present recommendations for centering decisions.

As noted above we present the results pertaining to centering in the context of examples. More formal presentation of these results can be found in Kreft, de Leeuw, and Aiken (1995) and Snijders and Berkhof (2008). Other expositions on centering can be found in Enders and Tofighi (2007); Hox (2002); Misangi, LePine, Algina, and Goeddeke (2006); Raudenbush and Bryk (2002); and Snijders and Bosker (1999).

As noted earlier, we present the major results about centering in the context of an example in which data are available for three variables: *MA*, *SES*, and a variable that indicates the school a participant attended. The *SES* data were taken from the High School and Beyond file supplied with HLM 6.0 but were transformed by adding five points to each student's *SES* score. This change was made in order to highlight the effect of centering on intercepts. The values of the dependent variable *MA* were simulated in order to ensure that the models we use to analyze the data are correct. The sample sizes are 7185 students in 160 schools. The model used to simulate the data for our first set of analyses was:

$$\mathcal{E}(MA_{ij}) = 15 + 2.2(SES_{ij}) + 3.8(\overline{SES}_j)$$

and

$$MA_{ij} = \mathcal{E}(MA_{ij}) + u_{0j} + u_{1j}SES_{ij} + \varepsilon_{ij}$$

where

- MA_{ij} is the mathematics achievement score for student i in school j
- SES_{ij} is the socioeconomic status score for student i in school j
- $\overline{SES}_j$ is the mean socioeconomic status score in school j
- $\mathcal{E}(MA_{ij})$ is the expected value of mathematics achievement conditional on SES_{ij} and $\overline{SES}_j$
- $u_{0j} + u_{1j}SES_{ij} + \varepsilon_{ij}$ is the residual (i.e., $MA_{ij} - \mathcal{E}(MA_{ij})$). The variables u_{0j} and u_{1j} were constants for students within school j but varied across schools. The variable ε_{ij} varied over students within a school.

TABLE 15.1

Descriptive Statistics for Math Achievement and Socioeconomic Status

Variable	***M***	***SD***
MA	45.22	7.00
SES	5.00	0.78

Note: $n = 7185$.

In the simulation, the coefficients for *SES* and school means *SES* were set to values similar to those obtained by analyzing the HSB data. The covariance matrix for the variables u_{0j} and u_{1j} was set as

$$\begin{bmatrix} 16.00 & \\ -2.50 & 0.50 \end{bmatrix}$$

and the variance for ε_{ij} was taken as 37. The variances and covariances above are similar to the values obtained by analyzing the HSB data. Descriptive statistics for the data are presented in Table 15.1.

15.2 SIMPLE LINEAR REGRESSION MODELS

If the data were not multilevel an appropriate analysis would be to use ordinary least squares (OLS) to estimate the simple linear regression model with *MA* as the dependent variable and *SES* as the independent variable:

$$MA_{ij} = \beta_0 + \beta_T SES_{ij} + \varepsilon_{ij}.$$

In this model, the two βs are fixed constants to be estimated and ε_{ij} is a random quantity. Fixed constants, either in regression models or multilevel models, are also called *fixed effects*. Random quantities are called *random effects* and the variance (and covariances) of random effects are called *variance components*. The T subscript in β_T stands for total and indicates that the grouping structure was not taken into account in the analysis.

The OLS estimate of the slope is $\hat{\beta}_T = 3.327$ and its standard error is 0.098.

To examine the effect of centering in this simple model, the independent variable is taken as the deviation of SES from the grand mean $\overline{\overline{SES}}$:

$$MA_{ij} = \beta_0 + \beta_T\left(SES_{ij} - \overline{\overline{SES}}\right) + \varepsilon_{ij}.$$

From Table 15.1, $\overline{\overline{SES}} = 5.00$. The estimated slope is again $\hat{\beta}_T = 3.327$ with a standard error of 0.098. That is, in simple linear regression, grand mean centering does not affect the estimated slope or its standard error.

Although grand mean centering did not affect the estimated slope or its standard error, grand mean centering does affect the estimate of the intercept. When *SES* is used as the independent variable the intercept is 28.587 and when grand mean centered *SES* is used as the independent variable the intercept is 45.222. Why are these estimates so different? When *SES* is not centered the estimated regression equation is

$$\widehat{MA} = 28.587 + 3.327(SES).$$

By definition the intercept is the predicted value when the independent variable is zero. When *SES* is used as the independent variable, zero is an *SES* score that is well below the grand mean of *SES* (5.00). As the standard deviation for *SES* is .78, a *SES* of zero is about 6.4 standard deviations below the grand mean. Thus 28.587 is predicted MA corresponding to an *SES* that is 6.4 standard deviations below the grand mean. When grand mean centered *SES* is used as the independent variable, the estimated regression equation is

$$\widehat{MA} = 45.222 + 3.327\left(SES - \overline{\overline{SES}}\right).$$

Now the intercept, 45.222, is the predicted value of *MA* when $SES - \overline{\overline{SES}} = 0$; that is, when $SES = 5.00$. So the reason the two intercepts are different is because they are predicted values corresponding to different *SES* scores. Another way to see the relationship between the slopes and the intercepts in the noncentered and the centered models is by noting that

$$MA_{ij} = \beta_0 + \beta_T SES_{ij} + \varepsilon_{ij} = \left(\beta_0 + \beta_T \overline{\overline{SES}}\right) + \beta_T\left(SES_{ij} - \overline{\overline{SES}}\right) + \varepsilon_{ij}.$$

Thus, while the slope remains the same, the intercept in the centered model is $\beta_0 + \beta_T \overline{\overline{SES}}$ and its estimate is 28.585 + 3.327(5) = 45.222.

Is the difference in the intercepts of fundamental importance? No, because any estimate or hypothesis test that can be obtained by using one of the two simple linear regression models can also be obtained by using the other model. If we use the model for noncentered *SES*, the predicted MA score for a student with an *SES* exactly equal to the grand mean is

$$\widehat{MA} = 28.587 + 3.327(5.00) = 45.222$$

and is equal to intercept under grand mean centering. A score of zero on the *SES* scale is equal to a score of 0–5.00 = –5.00 on the grand mean centered *SES* scale. If we use the model for grand mean centered *SES* and calculate the predicted value at the score –5.00 we get

$$\widehat{MA} = 45.222 + 3.327(0 - 5.00) = 28.587$$

Thus, given the results from one model, the intercept for the other model can be obtained by algebra. Models are statistically equivalent if any estimate or hypothesis test that can be obtained by using one of the models can also be obtained by using the other models.

The OLS method makes the important assumption that residuals are independent. This assumption is violated when the data are multilevel in nature as illustrated in Figure 15.1 where an idealized diagram of the data is presented. The large ellipse represents the scatter plot of all the data and the slanted line is the OLS regression line for the model $MA_{ij} = \beta_0 + \beta_T SES_{ij} + \varepsilon_{ij}$. The smaller ellipses represent the scatter plots for six schools. A residual is the deviation of a student's *MA* score from the regression line. The scatter plots for Schools A and B are wholly above the regression line and the residuals for the schools are all positive. The scatter plots for Schools E and F are wholly below the regression line and the residuals for the schools are all negative. The scatter plots for these schools illustrate that the residuals are similar in size within schools. As a result the residuals are statistically dependent, violating an important assumption of the OLS estimation procedure.

15.3 THE FIXED EFFECTS MODEL

One solution to the problem of nonindependent residuals is to use the fixed effects model (see, for example, Baltagi, 2005; Frees, 2004; or Greene, 2007). As applied in the current situation, the fixed effects model specifies a different intercept for each of the 160 schools and a common slope for all schools. The model is

$$MA_{ij} = \beta_{0j} + \beta_W SES_{ij} + \varepsilon_{ij}.$$

The coefficient β_{0j} is the intercept in the j^{th} school. These 160 intercepts are fixed effects that can be estimated but often are not because they are of limited interest. Treating the intercepts as fixed effects is

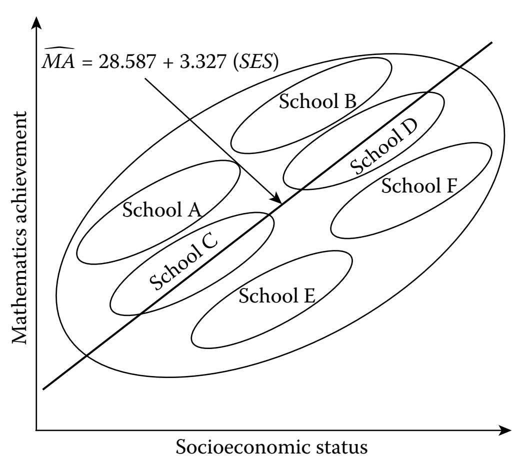

FIGURE 15.1
Depiction of the simple regression model.

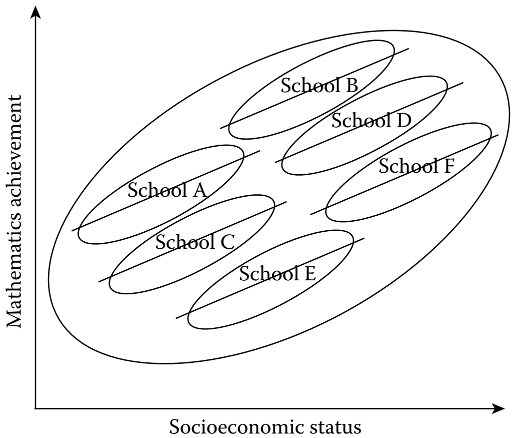

FIGURE 15.2
Depiction of the fixed effects model.

the reason the model is called a fixed effect model. The coefficient β_W is a fixed effect that will be estimated. The *W* subscript in β_W stands for "within." The fixed effects model is equivalent to the ANCOVA model and is estimated by OLS. As is well known, the slope in the ANCOVA model is assumed to be equal across schools and is a within-school slope; that is, the slope estimates the relationship between *MA* and *SES* within any one of the 160 schools. This is the reason for the *W* subscript in β_W.

The fixed effects model is depicted in Figure 15.2. Note, for example, the regression line for School A goes through the scatter plot for School A. The residual for a student within a school is defined as the deviation of that student's *MA* score from the regression line for the student's school. In School A, as in all of the other schools in Figure 15.2, there would be both positive and negative residuals and so the problem of nonindependent residuals is addressed. The coefficient $\hat{\beta}_W$ is equal to 2.181 with a standard error of .109. Note that $\hat{\beta}_W$ is approximately .65 the size of $\hat{\beta}_T$. For the purposes of comparing results, coefficients for the various models considered so far and several that will be considered subsequently are presented in Table 15.2.

15.4 THE RANDOM EFFECTS MODEL (RANDOM INTERCEPTS MODEL)

Another possible solution to the problem of nonindependent residuals is to use the random effects model (see, for example, Baltagi, 2005; Frees, 2004; Greene, 2007). The model can be written in the same form as the fixed effects model

$$MA_{ij} = \beta_{0j} + \beta_{SES}SES_{ij} + \varepsilon_{ij}.$$

but (a) β_{SES} is not necessarily equal to β_W, and (b) the β_{0j} are regarded as random quantities. Treating the intercepts as random quantities is the reason the model is called the random effects model. Since β_{0j} is a random variable, it can be expressed as $\beta_{0j} = \gamma_0 + u_{0j}$, where γ_0 is the expected value of β_{0j} and $u_{0j} = \gamma_0 - \beta_{0j}$ is the random error. Substituting $\beta_{0j} = \gamma_0 + u_{0j}$, the model can be written as

$$MA_{ij} = \gamma_0 + \gamma_{SES}SES_{ij} + u_{0j} + \varepsilon_{ij} \qquad (15.1)$$

where $\gamma_{SES} = \beta_{SES}$, and the fixed effects to be estimated are γ_0 and γ_{SES}. Both u_{0j} and

TABLE 15.2

Summary of Results for Selected Models

Model	Equation	Coefficient (Standard Error) for *SES*	Coefficient (Standard Error) for MEAN *SES*
Simple regression	$MA = \beta_0 + \beta_T SES + \varepsilon$	3.327 (0.098)	NA
Simple regression	$MA = \beta_0 + \beta_T\left(SES - \overline{\overline{SES}}\right) + \varepsilon$	3.327 (0.098)	NA
Fixed effects	$MA = \beta_{0j} + \beta_W(SES) + \varepsilon$	2.181 (0.109)	NA
Random intercepts	$MA = \gamma_0 + \gamma_{SES} SES + u_0 + \varepsilon$	2.373 (0.107)	NA
Random intercepts	$MA = \gamma_0 + \gamma_{SES}\left(SES - \overline{\overline{SES}}\right) + u_0 + \varepsilon$	2.373 (0.107)	NA
Random intercepts	$MA = \gamma_0 + \gamma_W\left(SES - \overline{SES}_j\right) + u_{0j} + \varepsilon$	2.181 (0.109)	NA
Intercepts as outcomes	$MA = \gamma_0 + \gamma_W(SES) + \gamma_C \overline{SES}_j + u_{0j} + \varepsilon$	2.181 (0.109)	4.186 (0.400)
Intercepts as outcomes	$MA = \gamma_0 + \gamma_W\left(SES - \overline{\overline{SES}}\right) + \gamma_C \overline{SES}_j + u_{0j} + \varepsilon$	2.181 (0.109)	4.186 (0.400)
Intercepts as outcomes	$MA = \gamma_0 + \gamma_W\left(SES - \overline{SES}_j\right) + \gamma_B \overline{SES}_j + u_{0j} + \varepsilon$	2.181 (0.109)	6.367 (0.385)

ε_{ij} are random effects. The variances of u_{0j} and ε_{ij} are the variance components. In the random effects model, the variance of u_{0j} is also the variance of the school specific intercept β_{0j}. The estimated coefficient for *SES* is $\hat{\gamma}_{SES} = 2.373$.

The term random effects model comes from the econometric literature. In the multilevel modeling literature, the random effects model is often called the random intercepts model. The model is depicted in Figure 15.3. As for the simple linear regression model, a single regression line for all schools is estimated. A residual is the deviation of a student's MA score from the regression line. Just as in the depiction for the simple linear regression model in Figure 15.1, the scatter plots for Schools A and B are wholly above the regression line and the residuals for the schools are all positive and the scatter plots for Schools E and F are wholly below the regression line and the residuals for the schools are all negative. Thus it would seem that the random effects model has the same problem as the simple linear regression model. However, in the random intercepts model the residual has two components $u_{0j} + \varepsilon_{ij}$ whereas the residual in the simple regression model has only one component. Because u_{0j} is equal for all students within a school, the inclusion of u_{0j} in the random intercepts model accounts for the similarity of residuals within schools and therefore the random intercepts model is more appropriate for the data than is the simple linear regression model.

Just as in the simple linear regression model, with the random intercepts model *SES* can be grand mean centered:

$$MA_{ij} = \gamma_0 + \gamma_{SES}\left(SES_{ij} - \overline{\overline{SES}}\right) + u_{0j} + \varepsilon_{ij}$$

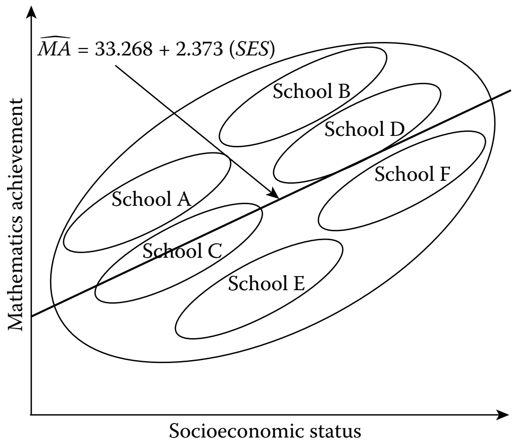

FIGURE 15.3
Depiction of the random effects model.

Grand mean centering does not affect $\hat{\gamma}_{SES}$ because

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_{SES} SES_{ij} + u_{0j} + \varepsilon_{ij} \\ &= \left(\gamma_0 + \gamma_{SES}\overline{\overline{SES}}\right) + \gamma_{SES}\left(SES_{ij} - \overline{\overline{SES}}\right). \\ &\quad + u_{0j} + \varepsilon_{ij}. \end{aligned}$$

As can be seen from the preceding equation, the intercepts when *SES* is not centered and when *SES* is grand mean centered are different and for the same reason presented in connection with the simple linear regression model. Because the two versions of the model are statistically equivalent, the intercept for each model can be calculated from the results for the other model. Similarly, while the estimated variance component for u_{0j}, which is the estimated variance for β_{0j}, will not be the same for the two models, the variance component for u_{0j} for one model can be obtained from the results for the other. These statements apply to all statistically equivalent models considered in this section and in the next section, and will not be repeated in each case.

15.5 INTERCEPTS AS OUTCOMES MODEL

The random intercepts model presented in Equation 15.1 assumes that the random effects (i.e., u_{0j} and r_{ij}) are uncorrelated with *SES*. Because $\beta_{0j} = \gamma_0 + u_{0j}$, the random intercepts model also assumes that β_{0j} is uncorrelated with *SES*. You can envision investigating this assumption by considering a plot with *SES* on the horizontal axis and β_{0j} on the vertical axis. If the assumption is met, the scatter plot would have a zero slope. Furthermore, if the assumption is met, a scatter plot with β_{0j} on the vertical axis and school mean *SES* $\left(\text{i.e., } \overline{SES}_j\right)$ on the horizontal axis would also have a zero slope. That is, assuming β_{0j} is uncorrelated with *SES* implies the assumption that β_{0j} is uncorrelated with $\overline{SES}_j$.

We can write the random intercepts model as a two-level model

Level-1:

$$MA_{ij} = \beta_{0j} + \beta_{SES} SES_{ij} + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + u_{0j}$$

$$\beta_{SES} = \gamma_{SES}.$$

Since β_{SES} is fixed, it does not have an error component; changing the notation from β_{SES} to $\beta_{SES} = \gamma_{SES}$ is merely for the purpose of conformity. Substituting the level-2 models into the level-1 model shows that the two-level model yields the random intercepts model in Equation 15.1. If we think the data violate the assumption that β_{0j} is uncorrelated with $\overline{SES}_j$ or wish to test the assumption, we can add school mean SES to the model for β_{0j}:
Level-1:

$$MA_{ij} = \beta_{0j} + \beta_{SES}SES_{ij} + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_C\overline{SES}_j + u_{0j}$$

$$\beta_{SES} = \gamma_W,$$

and our model becomes

$$MA_{ij} = \gamma_0 + \gamma_W\left(SES_{ij}\right) + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij}. \qquad (15.2)$$

This is an example of an intercepts as outcomes model.

In Equation 15.2 a noncentered level-1 variable and the mean of the level-1 variable are included in the model. The mean is called a contextual or compositional variable (Raudenbush, 1989) because it is a measure of the context of the group. The coefficient for the contextual variable is called a context coefficient and the subscript C on γ_C denotes context. Note that with the introduction of $\overline{SES}_j$ as an independent variable in the model, the subscript to the coefficient for SES has been changed from SES to W, signaling that including $\overline{SES}_j$ in the model changes the interpretation of the coefficient for SES. Estimates of the parameters are $\hat{\gamma}_W = 2.181$ and $\hat{\gamma}_C = 4.186$. (The standard errors are 0.109 and 0.400, respectively.) Note that the coefficient for SES is exactly the same in the fixed effects model and in the intercepts as outcomes model.

An alternative model using grand mean centering for SES is

$$MA_{ij} = \gamma_0 + \gamma_W\left(SES_{ij} - \overline{\overline{SES}}\right) + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij}.$$

Since

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_W\left(SES_{ij}\right) + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij} \\ &= \left(\gamma_0 + \gamma_W\overline{\overline{SES}}\right) + \gamma_W\left(SES_{ij} - \overline{\overline{SES}}\right) \\ &\quad + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij}, \end{aligned}$$

the grand mean centered model is statistically equivalent to the model in Equation 15.2 and the estimates of γ_W and γ_C are the same for the two models. The intercept in the grand mean centered model is, however, $\gamma_0 + \gamma_W\overline{\overline{SES}}$.

15.6 THE RANDOM INTERCEPTS MODEL WITH GROUP MEAN CENTERING

Recall that β_{0j} in the random intercepts model (see Equation 15.1) is assumed to be uncorrelated with SES and this assumption implies that β_{0j} is uncorrelated with school mean SES. Estimating the intercepts as outcomes model (see Equation 15.2) is one way to address this implication. Another way to address the assumption is by group (school) mean centering SES in the random intercepts model. To school mean center SES, we replace SES in

Equation 15.1 by $SES_{ij} - \overline{SES}_j$. The assumption now is that β_{0j} is uncorrelated with the school means of $SES_{ij} - \overline{SES}_j$. Because the school mean for $SES_{ij} - \overline{SES}_j$ must be zero for each school, the correlation between β_{0j} and the school-mean-centered variable will be zero. Thus by school mean centering *SES* we have ensured that the assumption is met.

The random intercepts model in which *SES* is group mean centered is

$$MA_{ij} = \gamma_0 + \gamma_W\left(SES_{ij} - \overline{SES}_j\right) + u_{0j} + \varepsilon_{ij}.$$

The estimated coefficient for *SES* is $\hat{\gamma}_W = 2.181$ with standard error 0.109 exactly the same as for the fixed effects model and the intercepts as outcomes model. However, $\hat{\gamma}_W$ is different than $\hat{\gamma}_{SES}$ that was obtained by using the random intercepts model with noncentered *SES* or with grand mean centered *SES*. This is because while the uncentered model is

$$MA_{ij} = \gamma_0 + \gamma_{SES}SES_{ij} + u_{0j} + \varepsilon_{ij},$$

the group mean centered model is

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_W\left(SES_{ij} - \overline{SES}_j\right) + u_{0j} + \varepsilon_{ij} \\ &= \gamma_0 + \gamma_W SES_{ij} + (-\gamma_W)\overline{SES}_j + u_{0j} + \varepsilon_{ij}. \end{aligned}$$

The group mean centered model is equivalent to a model that has two independent variables *SES* and $\overline{SES}_j$ and, in which, the coefficients for the two variables have opposite signs that are equal in absolute value. Because of this, the uncentered model is not statistically equivalent to the group mean centered model. Hence, for the random intercepts model, while the uncentered and the grand mean centered models are statistically equivalent producing identical slope parameters (and equivalent intercepts), these two models are not statistically equivalent to the random intercepts model with school mean centering.

15.7 THE INTERCEPTS AS OUTCOMES MODEL WITH GROUP MEAN CENTERING

As an alternative to the intercepts as outcomes model without centering or with grand mean centering, a model with school mean centering can be used:

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_W\left(SES_{ij} - \overline{SES}_j\right) + \gamma_B\left(\overline{SES}_j\right) \\ &\quad + u_{0j} + \varepsilon_{ij}. \end{aligned} \tag{15.3}$$

This model is statistically equivalent to the uncentered or the grand mean centered intercepts as outcome model. In comparing the uncentered model with the group mean centered model in Equation 15.3, we see that

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_W\left(SES_{ij} - \overline{SES}_j\right) + \gamma_B\left(\overline{SES}_j\right) \\ &\quad + u_{0j} + \varepsilon_{ij} \\ &= \gamma_0 + \gamma_W SES_{ij} + (\gamma_B - \gamma_W)\left(\overline{SES}_j\right) \\ &\quad + u_{0j} + \varepsilon_{ij}. \end{aligned}$$

Thus the coefficient of the group centered *SES* is the same as the coefficient of uncentered *SES*; the coefficient of $\overline{SES}$ in the group centered model is equal to $\gamma_B - \gamma_W$ in the uncentered model. As in the other two intercepts as outcomes models, the coefficient for *SES* is the within-school coefficient and is estimated to be $\hat{\gamma}_W = 2.181$ just as for the other two models. The coefficient for school mean *SES*, however, is not the context effect (γ_C). Instead it is γ_B, the between groups coefficient.

To help understand the difference between γ_C and γ_B consider the means as outcomes model

$$MA_{ij} = \gamma_0 + \gamma_B(\overline{SES}_j) + u_{0j} + \varepsilon_{ij}. \quad (15.4)$$

In this model the coefficient for the *SES* variable is a slope relating school specific *MA* means to school specific *SES* means. This is a between school relationship and therefore the coefficient for mean *SES* is denoted by γ_B; for the current example $\hat{\gamma}_B = 6.365$. Similarly the coefficient for mean *SES* in the intercepts as outcomes model with group mean centering is a between-school relationship; the coefficient for mean *SES* is also denoted by γ_B and the estimated coefficient is $\hat{\gamma}_B = 6.367$. (The estimates of γ_B in the means as outcomes model and the intercepts as outcomes model are not necessarily equal, but will be similar.) Earlier we found that $\hat{\gamma}_W = 2.181$ and $\hat{\gamma}_C = 4.186$. We can see that $\hat{\gamma}_C$ is equal to $\hat{\gamma}_B - \hat{\gamma}_W = 4.186$. Moreover, as was seen earlier, $\gamma_C = \gamma_B - \gamma_W$. That is, the context effect is simply the difference between the between-school effect and the within-school effect.

Recall that the models that include γ_C (i.e., the context effect) are

$$MA_{ij} = \gamma_0 + \gamma_W(SES_{ij}) + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij},$$

and

$$MA_{ij} = \gamma_0 + \gamma_W(SES_{ij} - \overline{\overline{SES}}) + \gamma_C\overline{SES}_j + u_{0j} + \varepsilon_{ij}.$$

The coefficient γ_C asks whether including school mean *SES* in the model is necessary if individual *SES* (in its non centered or grand mean centered forms) is also in the model. If $\gamma_C = 0$, it is not necessary to include school mean *SES* in the model and one can use the random intercepts model

$$MA_{ij} = \gamma_0 + \gamma_{SES}(SES_{ij}) + u_{0j} + \varepsilon_{ij}.$$

Also if $\gamma_C = 0$, then $\gamma_B = \gamma_W$ and

$$MA_{ij} = \gamma_0 + \gamma_W(SES_{ij} - \overline{SES}_j) + \gamma_B(\overline{SES}_j) + u_{0j} + \varepsilon_{ij} \quad (15.5)$$

simplifies to

$$MA_{ij} = \gamma_0 + \gamma_{SES}(SES_{ij}) + u_{0j} + \varepsilon_{ij},$$

implying it is unnecessary to include school mean *SES* in Equation 15.5. It can be shown that when $\gamma_B = \gamma_W$, $\hat{\gamma}_{SES}$ estimates the common coefficient and $\hat{\gamma}_{SES}$ has a smaller sampling variance than does $\hat{\gamma}_W$ or $\hat{\gamma}_B$.

According to Raudenbush and Bryk (2002), $\hat{\gamma}_{SES}$ is a weighted average of $\hat{\gamma}_B$ and $\hat{\gamma}_W$. This makes it clear that if $\gamma_B \neq \gamma_W$, then the random intercepts model without centering or with grand mean centering should not be used because it averages coefficients that provide information about two different aspects of the relationship between *MA* and *SES*.

The means as outcomes model in Equation 15.4 includes $\overline{SES}_j$ but does not include $SES - \overline{SES}_j$, whereas the intercepts as outcomes model with school mean centered *SES* in Equation 15.3 includes both variables. Nevertheless the estimate of the coefficient for $\overline{SES}_j$ (i.e., $\hat{\gamma}_B$) will be similar for the models. Why is the estimate of γ_B largely unaffected by the inclusion of $SES - \overline{SES}_j$ in the intercepts as outcomes model? Because $SES-\overline{SES}_j$ and $\overline{SES}_j$ are uncorrelated. If $SES-\overline{SES}_j$ is plotted against $\overline{SES}_j$, then at any point on the $\overline{SES}_j$ axis, the mean of $SES-\overline{SES}_j$ must be zero and so the plot will have a slope that is equal to zero. (This claim assumes that the participants who have a score on $SES-\overline{SES}_j$ are the same participants for whom $\overline{SES}_j$ was calculated.)

Rather than including $\overline{SES}_j$ as an independent variable we may want to include a level-2 variable that is not a mean of a level-1 independent variable. Following Susser (1994), we refer to a variable that is not a mean of a level-1 independent variable as an *integral* variable. Just as $SES-\overline{SES}_j$ and $\overline{SES}_j$ are uncorrelated, $SES-\overline{SES}_j$ will be uncorrelated with any level-2 integral variable. For example, suppose we want to study the relationship between the schools' disciplinary climate (*DC*) and *MA*. The correlation between *DC* and $SES-\overline{SES}_j$ will be zero. As a result the coefficient for *DC* will be similar in the following models (these analyses use the actual HSB data):

Means as outcomes model:

$$MA_{ij}=\gamma_0+\gamma_{DC}^{B}DC_j+u_{0j}+\varepsilon_{ij} \tag{15.6}$$

$$\widehat{MA}=12.59+(-1.49)DC$$

Intercepts as outcomes model:

$$\begin{aligned}MA_{ij}=\gamma_0+\gamma_W\left(SES-\overline{SES}_j\right)\\+\gamma_{DC}^{B}DC_j+u_{0j}+\varepsilon_{ij}\end{aligned} \tag{15.7}$$

$$\widehat{MA}=12.59+2.19\left(SES-\overline{SES}_j\right)+(-1.49)DC$$

where the superscript in γ_{DC}^{B} indicates a between-schools coefficient. So, including group mean centered *SES* in the model did not materially affect the estimate of γ_{DC}^{B}. In essence, by group mean centering *SES* in Equation 15.7 we have failed to control for *SES*. Alternatively we can say that we have controlled for the deviation of *SES* from school mean *SES*, but not for school mean *SES*. Regardless of how we describe the control, we do not know whether the relationship of *MA* to *DC* would be the same if we had a more complete control of *SES*. In our opinion models like that in Equation 15.7 should not be used unless the researcher wants to estimate the within-school relationship for the level-1 variables and a between-school relationship for the level-2 variables.

More complete control of *SES* can be achieved by including school mean *SES* in the model:

$$\begin{aligned}MA_{ij}=\gamma_0+\gamma_W\left(SES-\overline{SES}_j\right)+\gamma_{\overline{SES}}^{B}\overline{SES}_j\\+\gamma_{DC}DC_j+u_{0j}+\varepsilon_{ij},\end{aligned} \tag{15.8}$$

$$\begin{aligned}\widehat{MA}=12.63+2.19\left(SES-\overline{SES}_j\right)\\+3.11\overline{SES}+(-0.69)DC.\end{aligned}$$

The coefficient γ_{DC} does not have a *B* superscript because γ_{DC} is not a between-school coefficient when both $SES-\overline{SES}_j$ and $\overline{SES}_j$ are included in the model. The coefficient γ_{DC} does not have a *C* superscript because we reserve the term context effect for the coefficient of the mean of a level-1 variable (e.g., $\overline{SES}_j$) when the model also includes the level-1 variable and that variable is not centered or is grand mean centered. Either of the other two types of centering will provide a statistically equivalent model.

15.8 OLS ESTIMATION REVISITED

Just as school mean SES can be included in multilevel models, it can be included in a single-level regression model:

$$MA_{ij} = \beta_0 + \beta_W SES_{ij} + \beta_C \overline{SES}_j + \varepsilon_{ij},$$

$$MA_{ij} = \beta_0 + \beta_W \left(SES_{ij} - \overline{\overline{SES}}\right) + \beta_C \overline{SES}_j + \varepsilon_{ij},$$

and

$$MA_{ij} = \beta_0 + \beta_W \left(SES_{ij} - \overline{SES}_j\right) + \beta_B \overline{SES}_j + \varepsilon_{ij}.$$

For all three models $\hat{\beta}_W = \hat{\gamma}_W = 2.181$. The coefficient $\hat{\beta}_C$ will be similar to $\hat{\gamma}_C$ but not identical (unless the sample size is the same in all schools). Likewise, the coefficients $\hat{\gamma}_B$ and $\hat{\beta}_B$ will be similar (but will be identical if the sample size is the same in all schools).

If the regression models and multilevel models result in the same or similar coefficients, is it necessary to use the multilevel model? In the multilevel models the schools are viewed as having been sampled from a larger group of schools. If this reflects the researcher's point of view, then using $\hat{\gamma}_W$, and/or $\hat{\gamma}_B$ and/or $\hat{\gamma}_C$ will result in standard errors that correctly reflect this view. If the researcher's view is that the schools are a fixed set of schools of interest, then the standard errors for $\hat{\beta}_B$ and $\hat{\beta}_C$ will be correct. The standard error, however, for $\hat{\beta}_W$ will be incorrect. The correct standard error for $\hat{\beta}_W$ should be obtained by using the fixed effects model, but if the total sample size is large the two standard errors will be quite similar.

Finally it can be shown that $\hat{\beta}_T$ has a mathematical relationship to $\hat{\beta}_B$ and $\hat{\beta}_W$. Specifically, if R^2 is the proportion of variance due to the schools in a one-way ANOVA of the MA data, then

$$\hat{\beta}_T = R^2\hat{\beta}_B + (1 - R^2)\hat{\beta}_W.$$

(See, for example, Pedhazur, 1982). That is, $\hat{\beta}_T$ is a proportion of variance weighted average of $\hat{\beta}_B$ and $\hat{\beta}_W$ and will be an inappropriate coefficient if $\beta_W \neq \beta_B$.

15.9 SUMMARY AND RECOMMENDATIONS IN REGARD TO RANDOM INTERCEPTS AND INTERCEPTS AS OUTCOMES MODELS

15.9.1 Models without Level-2 Variables

The random intercepts model, without centering or with grand mean centering, entails the assumption that the school specific intercepts (β_{0j}) are uncorrelated with the independent variable. If this assumption is violated, or equivalently, if the within group (γ_W) and between group (γ_B) coefficients are not equal, the coefficients for the independent variables can be misleading in the random intercepts model without centering or with grand mean centering. If the researcher is only interested in within group coefficients, the random intercepts model with group mean centered independent variables should be used.

15.9.2 Models with Level-2 Variables

If the variables in the model are (a) level-1 variables, and (b) contextual level-2 variables that are means of the level-1 variables and if the researcher is interested in the within group coefficient (γ_W) and either the between group coefficient or the context coefficient (γ_C), any of the intercepts as outcomes models can be used. The models in which the independent variable is not centered or is grand mean centered may have greater utility because these models directly result in a test of $H_0: \gamma_C = 0$. However, if the independent variable is group mean centered $H_0: \gamma_C = 0$ can be tested by testing $H_0: \gamma_B - \gamma_W = 0$.

If the researcher is interested in an integral level-2 variable (e.g., disciplinary climate) in the model then the correct centering depends on the question the researcher plans to address. If the researcher is interested in the within-group effect of the level-1 variables and the between group effect of the level-2 variables, then the level-1 variables should be group mean centered. However, such models should be used with caution because the level-2 variables are investigated without complete control of the level-1 variables. In our opinion it will usually be preferable to include means of the level-1 variables in the model and then any of the three centering options can be used (see Equation 15.7, for example).

15.10 MODIFYING THE RANDOM INTERCEPTS AND INTERCEPTS AS OUTCOMES MODELS: ADDING A RANDOMLY VARYING SLOPE

The random intercepts model and the intercepts as outcomes model entail the assumption that the school specific slope is a constant across schools. Incorrectly making this assumption is not likely to have a strong impact on the coefficients in these models, but it can impact the standard errors and precludes estimation of the variance of the school specific slopes. Both the random intercepts model and the intercepts as outcomes models can be modified by specifying that the school specific slope varies across schools.

15.10.1 Random Regression Coefficients Models

By adding a randomly varying slope to the random intercepts model we obtain the random regression coefficients model. The version of this model without centering *SES* is Level-1:

$$MA_{ij} = \beta_{0j} + \beta_{1j} SES_{ij} + \varepsilon_{ij}.$$

Level-2:

$$\beta_{0j} = \gamma_0 + u_{0j}$$

$$\beta_{1j} = \gamma_{SES} + u_{1j}.$$

Combined

$$MA_{ij} = \gamma_0 + \gamma_{SES} SES_{ij} + u_{0j} + u_{1j} SES_{ij} + \varepsilon_{ij}. \tag{15.9}$$

As with the random intercepts models, there are two other variations of the random regression coefficients model, created by either grand mean centering or group mean centering the independent variable. When the independent variable is grand mean centered the model is

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_{SES}\left(SES_{ij} - \overline{\overline{SES}}\right) \\ &\quad + u_{0j} + u_{1j}\left(SES_{ij} - \overline{\overline{SES}}\right) + \varepsilon_{i} \end{aligned} \tag{15.10}$$

TABLE 15.3

Summary of Results for Random Regression Coefficients and Intercepts and Slopes as Outcomes Models

Model	Equation	Coefficient (Standard Error) for *SES*	Coefficient (Standard Error) for MEAN *SES*
Random regression coefficients	$MA = \gamma_0 + \gamma_{SES} SES + u_0 + u_1 SES + \varepsilon$	2.375 (0.127)	NA
Random regression coefficients	$MA = \gamma_0 + \gamma_{SES}(SES - \overline{\overline{SES}}) + u_0 + u_1(SES - \overline{\overline{SES}}) + \varepsilon$	2.375 (0.127)	NA
Random regression coefficients	$MA = \gamma_0 + \gamma_{SES}(SES - \overline{SES}_j) + u_0 + u_1(SES - \overline{SES}_j) + \varepsilon$	2.173 (0.127)	NA
Intercepts and slopes as outcomes	$MA = \gamma_0 + \gamma_W SES + \gamma_C \overline{SES}_j + u_0 + u_1 SES + \varepsilon$	2.186 (0.128)	4.115 (0.402)
Intercepts and slopes as outcomes	$MA = \gamma_0 + \gamma_W(SES - \overline{\overline{SES}}) + \gamma_C \overline{SES}_j + u_0 + u_1(SES - \overline{\overline{SES}}) + \varepsilon$	2.186 (0.128)	4.115 (0.402)
Intercepts and slopes as outcomes	$MA = \gamma_0 + \gamma_W(SES - \overline{SES}_j) + \gamma_B \overline{SES}_j + u_0 + u_1(SES - \overline{SES}_j) + \varepsilon$	2.183 (0.127)	6.295 (0.382)

The models in Equation 15.9 and Equation 15.10) are statistically equivalent. Because the intercepts in Equations 15.9 and 15.10 are defined for different values on the *SES* scale (see the discussion of the intercept for the simple regression model when the independent variable is not centered and when it is grand mean centered) the estimated intercepts (33.33 and 45.21), variance components for u_{0j} (29.85 and 5.68) and covariance for u_{0j} and u_{1j} (−4.18 and −0.64) are different for the two models. Nevertheless, the results for one of the models can be obtained from the results for the second. The estimated variance component for u_{1j} will be the same for the two models and is 0.71 for the example.

Each of the models given in Equations 15.9 and 15.10 assumes that the school specific intercept (β_{0j}) is independent of *SES*. If the data analyst does not want to make this assumption, the independent variable can be group mean centered. The resulting model is

$$\begin{aligned} MA_{ij} = \gamma_0 + \gamma_W(SES_{ij} - \overline{SES}_j) \\ + u_{0j} + u_{1j}(SES_{ij} - \overline{SES}_j) + \varepsilon, \end{aligned} \tag{15.11}$$

and is not statistically equivalent to the other two. Results for the three models are shown in Table 15.3. The coefficients for *SES* are equal for the first two models, but not for the third.

Type of centering affects the meaning of the intercepts and as a consequence the intercept (45.11) is different for Equation 15.11 than for Equation 15.9 or Equation 15.10. Similarly the variance component for u_{0j} (10.14) and covariance for u_{0j} and u_{1j} (−0.96) are different for Equation 15.11) than for Equations 15.9 and 15.10. For all three types of centering, the variance component for u_{1j} is the variance of the school specific slope (β_{1j}). Nevertheless the estimate is different for Equation 15.11 than for Equation 15.9 and Equation 15.10. For Equation 15.11 the estimate is 0.63. According to Raudenbush and Bryk (2002) group mean centering results in a better estimate of the variance component for u_{1j}.

Comparing results for the random intercepts (Table 15.2) and random regression coefficients (Table 15.3) models, we see that even with the same kind of centering, the coefficients vary somewhat across the two types of models. However, as in this example, the coefficients are typically fairly similar for the two model types.

15.10.2 Intercepts and Slopes as Outcomes Models

By adding a randomly varying slope to the intercepts as outcomes model we obtain an intercepts and slopes as outcomes model. The version of this model without centering *SES* is

Level-1:

$$MA_{ij} = \beta_{0j} + \beta_{1j} SES_{ij} + \varepsilon_{ij},$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_C \overline{SES}_j + u_{0j}$$

$$\beta_{1j} = \gamma_W + u_{0j},$$

Combined:

$$MA = \gamma_0 + \gamma_W SES + \gamma_C \overline{SES}_j + u_{0j} + u_{1j} SES + \varepsilon_{ij}. \quad (15.12)$$

There are two varieties of the intercepts and slopes as outcomes model. In the first variety, which is the subject of this section, the level-2 variable is a predictor of the intercept only. In the second variety, the subject of a subsequent section entitled "Models with Cross-level Interactions," the level-2 variable is a predictor of the slope.

Observations about the first variety of the intercepts and slopes as outcomes model are similar to the observations about random regression coefficients models. The model in Equation 15.12, with a noncentered level-1 independent variable and the model with a grand mean centered level-1 independent variable,

$$MA = \gamma_0 + \gamma_W \left(SES - \overline{\overline{SES}}\right) + \gamma_C \overline{SES}_j + u_{0j} + u_{1j} \left(SES - \overline{\overline{SES}}\right) + \varepsilon_{ij}, \quad (15.13)$$

are statistically equivalent. The coefficient for *SES* in these models is a within school coefficient and the coefficient for school means *SES* is a context coefficient. When the level-1 independent variable is group mean centered, the model is

$$MA_{ij} = \gamma_0 + \gamma_W \left(SES_{ij} - \overline{SES}_j\right) + \gamma_B \overline{SES}_j + u_{0j} + u_{1j} \left(SES_{ij} - \overline{SES}_j\right) + \varepsilon_{ij} \quad (15.14)$$

and is not statistically equivalent to the other two models. The fact that the three models are not statistically equivalent stands in contrast to the equivalence status of the three intercepts as outcomes model.

Recall that we defined statistically equivalent models as models for which any estimate or test statistic obtained by using one of the models can be obtained by using the other models. Kreft et al. (1995) classified models as equivalent in the fixed effects and/or equivalent in the variance components. Models are equivalent in the fixed effects if the population fixed effects for one model can be expressed as functions of the population fixed effects of the other model. Models are equivalent in the variance components if the population variance components for one model can be expressed as functions of the population variance components of the other model. Models that are statistically equivalent by our definition are equivalent by both of the criteria set forth by Kreft et al. (1995). The models in Equations 15.12, 15.13, and 15.14 are equivalent in the fixed effects. The model in Equation 15.14 is not equivalent to the other models in the variance components.

Despite the fact that γ_W is the same parameter for Equation 15.12 to Equation 15.14, the estimate of γ_W for Equation 15.14 is not necessarily the same as is the estimate of γ_W in Equations 15.12 and 15.13. This occurs because the models are not equivalent in the variance components and the variance components are used in estimating the fixed effects. In the present example $\hat{\gamma}_W$ for Equation 15.14 differs from $\hat{\gamma}_W$ in Equations 15.12 and 15.13 in the third decimal place. This is consistent with our experience that all three models will provide similar estimates of γ_W. With the intercepts as outcomes model, $\hat{\gamma}_C = \hat{\gamma}_B - \hat{\gamma}_W$. This is not necessarily true with the intercepts and slopes as outcomes model (see Table 15.3). Comparing results for the intercepts as outcomes (Table 15.2) and intercepts and slopes as outcomes (Table 15.3) models we see that even with the same kind of centering, the coefficients vary across the two types of models.

In Section 15.7 we discussed models in which a level-2 variable is an integral variable rather than a contextual variable. With the following exceptions, the discussion in Section 15.7 applies to the first variety of intercepts and slopes as outcomes model.

1. The difference in the estimates of γ_{DC}^B in the means as outcomes model (see Equation 15.7) and in the intercepts as outcomes model (see Equation 15.6) is likely to be smaller than the difference in the estimates of γ_{DC}^B in the means as outcomes and in intercepts and slopes as outcomes model:

$$\begin{aligned} MA_{ij} &= \gamma_0 + \gamma_W \left(SES - SES_j\right) + \gamma_{DC}^B DC_j \\ &\quad + u_{0j} + \left(SES - SES_j\right) u_{1j} + \varepsilon_{ij}. \end{aligned} \tag{15.15}$$

2. The model in Equation 15.15 is not statistically equivalent to the versions of this model in which the level-1 variable is not centered or it is grand mean centered, though it is equivalent in the fixed effects.

15.11 RECOMMENDATIONS: IN REGARD TO RANDOM REGRESSION COEFFICIENTS AND INTERCEPTS AND SLOPES AS OUTCOMES MODELS

When a researcher is only interested in the within group effect (γ_W), the random regression coefficients model with group mean centering should be used. In addition, according to Raudenbush and Bryk

(2002) using group mean centering provides a better estimate of the variance components for u_{1j}.

When the researcher is interested in the within group coefficient and the context coefficient (γ_C) the intercepts and slopes as outcomes model can be used. The independent variable should either be noncentered or grand mean centered. When the researcher is interested in the within group coefficient and the between group (γ_B) the intercepts and slopes as outcomes model with a group mean centered independent variable can be used.

15.12 MODELS WITH CROSS-LEVEL INTERACTIONS

As noted earlier, the model with *SES* and school mean *SES* as independent variables,

$$MA_{ij} = \gamma_0 + \gamma_W SES + \gamma_C \overline{SES}_j + u_{0j} + u_{1j}SES + \varepsilon_{ij},$$

is an example of the first variety of the intercepts and slopes as outcomes model. One assumption of this model is that the school specific slope β_{1j} is uncorrelated with $\overline{SES}_j$. To address this assumption the following intercepts and slopes as outcomes model, an example of the second variety, can be used:

Level-1:

$$MA_{ij} = \beta_{0j} + \beta_{1j}SES_{ij} + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_2 \overline{SES}_j + u_{0j}$$

$$\beta_{1j} = \gamma_1 + \gamma_3 \overline{SES}_j + u_{1j},$$

Combined:

$$MA_{ij} = \gamma_0 + \gamma_1 SES + \gamma_2 \overline{SES}_j + \gamma_3 (SES \times \overline{SES}_j) + u_{0j} + u_{1j}SES + \varepsilon_{ij}. \tag{15.16}$$

The inclusion of $SES \times \overline{SES}_j$ allows investigation of the cross-level interaction. Alternatives to Equation 15.16 replace *SES* by grand mean centered *SES* or by group mean centered *SES*. As usual, the model with group mean centered *SES* is not statistically equivalent to the other two, which are statistically equivalent. In addition, as shown by Kreft et al. (1995) the model with group mean centered *SES* is not equivalent

TABLE 15.4

Summary of Results for Intercept and Slopes as Outcomes Model with a Cross-Level Interaction

Random Effect for the Slope	Centering for *SES*	Coefficient (Standard Error) for *SES* $(\hat{\gamma}_1)$	Coefficient (Standard Error) for Mean *SES* $(\hat{\gamma}_2)$	Coefficient (Standard Error) for Interaction $(\hat{\gamma}_3)$
Yes	None	2.074 (1.568)	4.056 (1.592)	1.524 (0.314)
	Grand Mean	2.074 (1.568)	11.674 (0.378)	1.524 (0.314)
	Group Mean	1.752 (1.660)	21.157 (0.361)	1.587 (0.332)
No	None	1.943 (1.285)	3.937 (1.381)	1.547 (0.257)
	Grand Mean	1.943 (1.285)	11.673 (0.376)	1.547 (0.257)
	Group Mean	1.672 (1.378)	21.157 (0.361)	1.602 (0.276)

TABLE 15.5

Summary of Simulation Results for Intercept and Slopes as Outcomes Model with a Cross-Level Interaction

Centering for *SES*	Coefficient (Standard Deviation) for *SES* ($\hat{\gamma}_1$)	Coefficient (Standard Deviation) for Mean *SES* ($\hat{\gamma}_2$)	Coefficient (Standard Deviation) for Interaction ($\hat{\gamma}_3$)
None	2.19	3.81	1.50
	(1.43)	(0.72)	(0.29)
Grand mean	2.19	11.32	1.50
	(1.43)	(0.42)	(0.29)
Group mean	2.20	20.84	1.50
	(1.52)	(0.42)	(0.30)

to the other two models in the fixed effects or in the variance components.

To generate data to compare the results of applying the three model variations, we used the following model:

$$\mathcal{E}(MA_{ij}) = 15 + 2.2(SES_{ij}) + 3.8(\overline{SES}_j) + 1.5(SES_{ij} \times \overline{SES}_j)$$

and

$$MA_{ij} = \mathcal{E}(MA_{ij}) + u_{0j} + u_{1j}SES_{ij} + \varepsilon_{ij}.$$

The covariance matrix for the u_{0j} and u_{1j} was again

$$\begin{bmatrix} 16.00 & \\ -2.50 & 0.50 \end{bmatrix}$$

and the variance for ε_{ij} was 37.

The first three lines in the body of tables 15.4 contain results for the three models. Note that the estimate of the coefficient for the product term (γ_3) when *SES* is group mean centered is different than the estimate of γ_3 for the other types of centering and is less similar to 1.5, which is the true value of γ_3. To determine if this finding is specific to this particular set of simulated data, we replicated the simulation 1000 times. The results are reported in Table 15.5 and indicate little if any difference among the three centering methods in the parameter being estimated by $\hat{\gamma}_3$. We also calculated the standard deviation of the difference between the estimate of γ_3 under group mean centering and the estimate under either of the other types of centering. These latter two estimates must be equal. The standard deviation was .09, indicating that estimates of γ_3 were fairly similar across centering methods in most replications.

In our example, centering had a minimal effect on the estimate of γ_3. However, the type of centering used in Equation 15.16 can affect the estimate of γ_3 and when it does, it typically means that an independent variable has been omitted from the model. For example, if the following model is correct for the data

$$\mathcal{E}(MA_{ij}) = \gamma_0 + \gamma_1 SES_{ij} + \gamma_2 \overline{SES}_j + \gamma_4 DC_j + \gamma_5 \overline{SES}_j \times DC_j, \tag{15.17}$$

and if $\overline{SES}_j \times DC_j$ and $\overline{SES}_j(SES_{ij} - \overline{\overline{SES}})$ are correlated, then in

$$\mathcal{E}(MA_{ij}) = \gamma_0 + \gamma_1(SES_{ij} - \overline{\overline{SES}}) + \gamma_2\overline{SES}_j + \gamma_3\overline{SES}_j(SES_{ij} - \overline{\overline{SES}}), \tag{15.18}$$

γ_3 will not be equal to zero and spurious evidence for the cross-level interaction of school mean *SES* and individual *SES* can emerge. The same conclusion holds for the version of Equation 15.17 in which *SES* is not centered. However, if *SES* is group mean centered the correlation between $\overline{SES}_j \times DC_j$ and $\overline{SES}_j(SES_{ij} - \overline{SES}_j)$ must be zero and, assuming Equation 15.17 is the correct model for the data, γ_3 will be equal to zero when the version of Equation 15.18 with a group mean centered *SES* is used.

As an alternative to using Equation 15.18, the following equation can be used

$$\mathcal{E}(MA_{ij}) = \gamma_0 + \gamma_1 SES_{ij} + \gamma_2\overline{SES}_j + \gamma_3\overline{SES}_j \times SES_{ij} + \gamma_4 DC_j + \gamma_5\overline{SES}_j \times DC_j.$$

Then assuming Equation 15.17 is the correct equation, spurious evidence for the cross-level interaction of school mean *SES* and individual *SES* should not emerge. That is, if the researcher knows what integral variable (i.e., DC_j) is likely to interact with the contextual variable (i.e., $\overline{\text{SES}}_j$), then the product term $\overline{SES}_j \times DC_j$ can be included in the model and there should not be spurious evidence that $\gamma_3 \neq 0$. This will be true regardless of which type of centering is used for *X*.

Similarly, Raundenbush (1989; see also, Enders & Tofighi, 2007; & Hoffman and Gavin, 1998) pointed out that if Equation 15.17 is correct, then using

$$\mathcal{E}(MA_{ij}) = \gamma_0 + \gamma_1(SES_{ij} - \overline{\overline{SES}}) + \gamma_4 DC_j + \gamma_6(SES_{ij} - \overline{\overline{SES}})DC_j, \tag{15.19}$$

could result in spurious evidence that $\gamma_6 \neq 0$ because $\overline{SES}_j \times DC_j$ and $(SES_{ij} - \overline{\overline{SES}})DC_j$ will be correlated. The same conclusion applies to the version of Equation 15.19 in which *SES* is not centered. This problem can be overcome by using group mean centering. In addition, as pointed out by Enders and Tofighi (2007), if $H_0\colon \gamma_6 = 0$ is tested by using

$$\mathcal{E}(MA_{ij}) = \gamma_0 + \gamma_1(SES_{ij} - \overline{SES}_j) + \gamma_2\overline{SES}_j + \gamma_4 DC_j + \gamma_5\overline{SES}_j \times DC_j + \gamma_6(SES_{ij} - \overline{SES}_j)DC_j,$$

then, assuming Equation 15.17 is correct, spurious test results for the test of $H_0 : \gamma_6 = 0$ should not emerge. The same conclusion holds regardless of the centering of *SES*. The general principal in these two examples is that in multilevel models as in regression models, it is essential to avoid omitting important variables from the model. Raudenbush and Bryk (2002, Chapter 10) describe a procedure for detecting omitted level-2 variables.

In Table 15.4, the estimate of the coefficient for the SES variable (γ_1) is quite different when *SES* is group mean centered than it is for the other two types of centering, but results in Table 15.5, for the 1000 simulated data sets, indicate little if any difference among the three centering methods in the parameters estimated by $\hat{\gamma}_1$. The standard deviation of the difference between the estimates of γ_1 under group mean centering and the estimate of γ_1 under either of the other type of centering was .50, indicating that the estimate of γ_1 under group mean centering could be quite different than the estimate under the other types of centering. This is consistent with the results in Table 15.4.

In Table 15.4, the estimate of the coefficient for school mean *SES* (γ_2) varies dramatically across the three methods of centering. We can understand these differences by considering the formula for the estimated simple slope for school mean *SES*. The general expression for a simple slope is (a) the coefficient for the variable of interest, which is school mean *SES* in the present example and is equal to $\hat{\gamma}_2$, plus (b) the coefficient of the product term, which is $\hat{\gamma}_3$, multiplied by the other variable in the product term, which will be one of the forms of *SES*. Under group mean centering, the formula for the simple slope for $\overline{SES}_j$ is $\hat{\gamma}_2+\hat{\gamma}_3\left(SES-\overline{SES}_j\right)$. Substituting a value for $\left(SES-\overline{SES}_j\right)$ in the formula provides an estimate of the slope for $\overline{SES}_j$ among students at that level of $\left(SES-\overline{SES}_j\right)$. For example, if we focus attention on students with an *SES* that is two points above the mean *SES* for their school we find 21.157 + 1.587(2) = 24.331, the simple slope for school mean *SES* among students whose *SES* is two points above the mean $\underline{SES}$ for their school. Based on $\hat{\gamma}_2+\hat{\gamma}_3\left(SES-\overline{SES}_j\right)$, $\hat{\gamma}_2$ is obtained by substituting zero for $\left(SES-\overline{SES}_j\right)$ and is the simple slope for school mean *SES* among students whose *SES* is equal to their school mean *SES*. Thus $\hat{\gamma}_2 = 21.157$ is the simple slope for school mean *SES* among students whose *SES* is equal to their school mean *SES*.

When *SES* is not centered, the simple slope for school mean *SES* is $\hat{\gamma}_2+\hat{\gamma}_3 SES$. Using $\hat{\gamma}_2+\hat{\gamma}_3 SES$, $\hat{\gamma}_2$ is obtained by substituting zero for *SES* and is the slope for school mean *SES* among students whose *SES* is 0. Thus $\hat{\gamma}_2 = 4.056$ is the simple slope for school mean *SES* among students whose *SES* is 0. To obtain the simple slope formula for grand mean centering, replace *SES* in $\hat{\gamma}_2+\hat{\gamma}_3 SES$, by $\left(SES-\overline{\overline{SES}}\right)$ to obtain $\hat{\gamma}_2+\hat{\gamma}_3\left(SES-\overline{\overline{SES}}\right)$. Therefore, under grand mean centering $\hat{\gamma}_2 = 11.674$ is the simple slope for school mean *SES* among students whose $SES-\overline{\overline{SES}}$ is equal to zero or, equivalently, whose *SES* is at the grand mean. These considerations show that $\hat{\gamma}_2$ varies across the centering because $\hat{\gamma}_2$ is a simple slope for school mean *SES* when the *SES* independent variable is zero and the meaning of zero on the *SES* independent variable varies across the three centering methods.

Recall that $\hat{\gamma}_2 = 11.674$ in the model with grand mean centered *SES* and that $\hat{\gamma}_2$ is the simple slope for school mean *SES* among students whose *SES* is at the grand mean. The simple slope for mean *SES* in the model in which SES is not centered is $\hat{\gamma}_2+\hat{\gamma}_3 SES$. We can use this expression to find the simple slope for school mean *SES* among students whose *SES* is at the grand mean by substituting 5 for *SES*. We find 4.056 + 1.524(5) = 11.676, which is within rounding error of the result obtained by using the model in which *SES* was grand mean centered. This illustrates that when models are equivalent, any estimate that can be obtained from one model can also be obtained from the other model.

Comparison of the formula for the simple slope for school mean *SES* under group mean centering $\left[\hat{\gamma}_2+\hat{\gamma}_3\left(SES-\overline{SES}_j\right)\right]$ to either the formula when there is no centering $[\hat{\gamma}_2+\hat{\gamma}_3 SES]$ or the formula when there is grand mean centering $\left[\hat{\gamma}_2+\hat{\gamma}_3\left(SES-\overline{\overline{SES}}\right)\right]$

shows that the nature of the simple slope is quite different under group mean centering. Specifically $\hat{\gamma}_2 + \hat{\gamma}_3(SES - \overline{SES}_j)$ estimates the effect of school mean *SES* for students who are a particular distance from their school mean *SES*. For example if $(SES - \overline{SES}_j) = 1$ then $\hat{\gamma}_2 + \hat{\gamma}_3(SES - \overline{SES}_j) = 21.157 + 1.587(1) = 22.744$ and tells us that among students whose *SES* is one point above the school mean *SES*, the effect of school mean *SES* is about 23 points. According to the model, this implication holds regardless of the school's mean *SES*. So if we focus attention on students with an SES of 4.5 in schools with an average *SES* of 3.5 (a low *SES*) or on students with an SES of 7.5 in a school with an average *SES* of 6.5 (a high *SES*), the effect of school mean *SES* will be 23 points. By contrast, according to either of the other models the simple slope for school mean *SES* depends on the student's actual *SES*.

Cross-level interactions can also be investigated by using intercepts and slopes as outcomes models without a random effect for the slope. If *SES* is not centered the model is

$$
\begin{aligned}
MA_{ij} &= \gamma_0 + \gamma_1 SES + \gamma_2 \overline{SES}_j \\
&\quad + \gamma_3 (SES \times \overline{SES}_j) + u_{0j} + \varepsilon_{ij}
\end{aligned}
$$

Results for this model, as well as results for variations on this model obtained by grand mean or group mean centering, are presented in the last three lines in the body of table 15.4. The effect of deleting the u_{1j} term on the coefficients is fairly small. But it is known that failing to include a required random effect tends to result in standard errors that underestimate the sampling variability in the estimates. Consistent with this result, the standard errors for the estimates tend to be smaller when the u_{1j} term is not included.

15.13 RECOMMENDATIONS FOR MODELS WITH CROSS-LEVEL INTERACTIONS

One issue is whether or not to include the u_{ij} term in the model. Whenever possible the u_{ij} term should be included. Another issue is the type of centering for the level-1 independent variables. In our simulation, centering had little effect on the interaction effect. These results no doubt reflect the fact that we used the same model to simulate the data and to analyze the data. Nevertheless, our experience is that centering often has little effect on the interaction coefficient. When centering has a strong effect on the estimate of the interaction parameter, it typically means that an independent variable has been omitted from the model. If centering does not have a strong impact on the cross-level interaction (i.e., the interaction is significant both when group mean centering is used and when it is not or is not significant in both cases) then we recommend against group mean centering unless the researcher wants to investigate the simple effect for the level-2 variable defined by how far removed the participant is from the group mean on the level-1 variable and not on the participant's actual score on the level-1 variable. If centering does have a strong impact on the cross-level interaction (i.e., the interaction is not significant when group mean centering is used but is significant with the other two types of centering) then we recommend excluding the cross-level-interaction from the model.

15.14 CENTERING LEVEL-2 VARIABLES

When a level-2 variable is centered, the new model is statistically equivalent to the original model. Therefore any estimate or hypothesis test that can be obtained by using the original model can also be obtained by using the revised model. Consequently centering level-2 variables is a less important issue than is centering level-1 variables. To illustrate let us consider Equation 15.16 and the model obtained by grand mean centering $\overline{SES}_j$:

$$MA_{ij} = \gamma_0 + \gamma_1 SES + \gamma_2\left(\overline{SES}_j - \overline{\overline{SES}}\right) + \gamma_3\left(SES \times \left[\overline{SES}_j - \overline{\overline{SES}}\right]\right) + u_{0j} + u_{1j}SES + \varepsilon_{ij}.$$

Results for Equation 15.16, originally reported in Table 15.4, are $\hat{\gamma}_1 = 2.074$, $\hat{\gamma}_2 = 4.056$, and $\hat{\gamma}_3 = 1.524$. After grand mean centering $\hat{\gamma}_1 = 9.692$, $\hat{\gamma}_2 = 4.056$, and $\hat{\gamma}_3 = 1.524$. Grand mean centering did not affect $\hat{\gamma}_2$ or $\hat{\gamma}_3$ because their interpretation is independent of level-2 centering. Grand mean centering did affect $\hat{\gamma}_1$ because its interpretation is affected by level-2 centering:

- When $\overline{SES}_j$ is the level-2 variable $\hat{\gamma}_1$ is the simple slope for *SES* when $\overline{SES}_j$ is zero
- When $\overline{SES}_j - \overline{\overline{SES}}$ is the level-2 variable $\hat{\gamma}_1$ is the simple slope for *SES* when $\overline{SES}_j$ is equal to the grand *SES* mean

Similarly when the level-1 variable is centered (either grand mean or group mean) in the cross-level interaction model, centering the level-2 variable only affects the estimate of γ_1. In the multilevel models considered in this chapter that do not include cross-level interactions, centering the level-2 variable only affects the intercept.

15.15 MODELS WITH REPEATED MEASURES

As an example of a repeated measures design, consider the data provided with HLM 6.0 in which adolescents were asked about their attitudes toward deviant behavior (*ATT*) and exposure to deviant peers (*EXP*) each year from age 11 to age 15. (For the sake of simplicity we used a version of this data set in which all participants with missing data were eliminated from the data file.) Henceforth we refer to data like those in the *MA* and *SES* example as between-subjects data because all of the variables vary between-subjects. We refer to data like those in the *ATT* and *EXP* example as mixed data because the variables vary within-subjects (over time) and between-subjects. The same issues we developed in the context of the *MA* and *SES* example could be developed in the context of this new example. Rather than repeating the developments we introduce some issues that apply primarily or uniquely in repeated measures designs.

Suppose the researchers are interested in the relationship of *ATT* to *EXP* and the researchers use the model

$$ATT_{ij} = \gamma_0 + \gamma_{EXP} EXP_{ij} + u_{0j} + \varepsilon_{ij}, \quad (15.20)$$

where ATT_{ij} is the attitude for person j at age i. By using this model the researchers run the risk of confounding the within person relationship between the two variables (how *ATT* changes over time with changes in *EXP*) with the between person relationship (how *ATT* scores vary across people in relation to variation in *EXP* scores across people.) Based on the developments for between-subjects data, the two relationships can be separated by using the model

$$ATT_{ij} = \gamma_0 + \gamma_W EXP_{ij} + \gamma_C \overline{EXP}_j + u_{0j} + \varepsilon_{ij}, \quad (15.21)$$

where $\overline{EXP}_j$ is the average *EXP* score for person *j* over the five ages. Separation of the two aspects of the relationship can also be achieved by using the other centering options. With group mean centering, the *EXP* independent variable becomes $EXP_{ij} - \overline{EXP}_j$ and is the deviation of the j^{th} person's *EXP* score at age *i* from that person's mean over the five ages. Thus, group mean centering in a repeated measures design is person mean centering. With grand mean centering, the *EXP* independent variable becomes $EXP_{ij} - \overline{\overline{EXP}}$ and is the deviation of the j^{th} person's *EXP* score at age *i* from the grand mean; that is, the mean over all ages and participants. Applying Equation 15.21 we obtain $\hat{\gamma}_0 = .324$, $\hat{\gamma}_W = .465$, and $\hat{\gamma}_C = .071$. Testing $H_0 : \gamma_C = 0$ tests the same hypothesis that is tested in Hausman's (1978) test to determine if a fixed or random effects model is appropriate for the data. If $H_0 : \gamma_C = 0$ is rejected then Equation 15.21 is preferred over Equation 15.20 and, as noted above, yields $\hat{\gamma}_W$ that is equal to the coefficient one would obtain by using the fixed effects model.

The purpose of including the person mean of *EXP* is to account for the possibility that u_{0j} is correlated with *EXP* in Equation 15.20 or equivalently that the within person intercept $\beta_{0j} = \gamma_0 + u_{0j}$ is correlated with *EXP*. There is another approach to accounting for this possibility that can be used with repeated measures data. The model is

$$ATT_{ij} = \gamma_0 + \gamma_W EXP_{ij} + u_{0j} + \varepsilon_{ij}, \quad (15.22)$$

but the estimation procedure allows for the possibility that u_{0j} is correlated with *EXP* at each occasion (Allison, 2005). Allowing for the possibility that u_{0j} is correlated with *EXP* at each occasion results in estimation of a within person coefficient. This procedure can be implemented in any structural equation modeling program. Appendix A presents an M*plus* program for implementing the method. The code—`a1 a2 a3 a4 a5 (3)`—restricts the variance of the residuals to be equal at all ages to maximize similarity of the results obtained by using Equation 15.21; the restriction is not necessary. The results are $\hat{\gamma}_0 = .324$ and $\hat{\gamma}_W = .465$. The estimate $\hat{\gamma}_W$ and its standard error are equal to the estimate $\hat{\beta}_W$ and its standard error that would be obtained by using the fixed effects model. A likelihood ratio test comparing the model that specifies u_{0j} is correlated with *EXP* at each occasion to one without the specification can be used to select between the two approaches. Without the restriction on the residual variances, the results are $\hat{\gamma}_0 = .319$ and $\gamma_W = .456$.

Estimation of Equation 15.22 allowing for the possibility that u_{0j} is correlated with *EXP* at each occasion can also be implemented in a hierarchical linear modeling program by writing the model as
Level-1:

$$ATT_{ij} = \beta_{0j} + \beta_1 EXP_{ij} + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_{01} EXP_{11j} + \cdots + \gamma_{05} EXP_{15j} + u_{oj}$$

and

$$\beta_1 = \gamma_W .$$

where $EXP_{11j},\ldots,EXP_{15j}$ are the exposure variables at the five ages. The level-2 equation is equivalent to Chamberlin's (1982, 1984) specification for the relationship of u_{0j} to the predictors in the model. The combined model is

$$ATT_{ij} = \gamma_0 + \gamma_W EXP_{ij} + \gamma_{01} EXP_{11j} + \cdots + \gamma_{05} EXP_{15j} + u_{0j} + \varepsilon_{ij}, \qquad (15.23)$$

Comparing Equations 15.23 and 15.21 shows that if $\gamma_{01} = \ldots = \gamma_{05} \equiv \gamma$ then Equation 15.23 simplifies to Equation 15.21 with $\gamma_C = 5\gamma$. To estimate Equation 15.23, we used the multivariate linear two-level model with heterogeneous residual variances in HLM 6.0. A screen shot of the program is presented in Appendix B. The results obtained were $\hat{\gamma}_0 = .302$ and $\gamma_W = .465$. The coefficient $\hat{\gamma}_0$ is different in Equation 15.23 and Equation 15.22, but the intercept obtained by using Equation 15.22 can be obtained by revising Equation 15.23 with each of $EXP_{11j},\ldots,EXP_{15j}$ centered around its mean. The hypothesis $H_0{:}\gamma_{01} = \ldots = \gamma_{05} = 0$ can be tested to determine if correlations between the person-specific intercepts and the exposure variable at each age are required in the model. The hypothesis tested in Hausman's test is a special case of $H_0{:}\gamma_{01} = \ldots = \gamma_{05} = 0$ because Hausman's approach assumes $\gamma_{01} = \ldots = \gamma_{05} = \gamma_C/5$.

Another approach to centering in repeated measures designs involves the concept of the cross-sectional and the longitudinal effects (see Diggle, Heagerty, Liang, & Zeger, 2002). Consider the two-level model

Level-1:

$$ATT_{ij} = \beta_{0j} + \beta_{1j}\left(EXP_{ij} - EXP_{11j}\right) + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_X EXP_{11j} + u_{0j}$$

and

$$\beta_{1j} = \gamma_L$$

where the subscript X denotes cross-sectional and the subscript L abbreviates longitudinal. Rather than group mean or grand mean centering *EXP*, the variable $EXP_{ij}–EXP_{11j}$ expresses *EXP* as a deviation from the value of *EXP* at age 11. The combined model is

$$ATT_{ij} = \gamma_0 + \gamma_X EXP_{11j} + \gamma_L\left(EXP_{ij} - EXP_{11j}\right) + u_{0j} + \varepsilon_{ij}.$$

In the level-1 equation, *EXP* is expressed as a deviation from *EXP* at age 11. As a consequence of this new type of centering, the intercept β_{0j} is expected *ATT* when *EXP* is equal to its value at age 11. We can think of β_{0j} as model-implied attitude for the adolescent j at age 11. An adolescent with a high β_{0j} will tend to have a high *ATT* at age 11 and an adolescent with a low β_{0j} will tend to have a low *ATT* at age 11. The coefficient γ_X is the cross sectional effect and measures the effect of *EXP* at age 11 on *ATT* at age 11. The variable $EXP_{ij}– EXP_{11j}$ measures change over time in *EXP*. Thus γ_L measures the effect of changes in *EXP* on *ATT*. The results are $\hat{\gamma}_X = .409$, implying that at age 11 adolescents with more exposure to deviant peers have a more positive attitude toward deviance, and $\hat{\gamma}_L = .504$, implying that as exposure to deviant peers increases (decreases) overtime, attitude toward deviance increases (decreases).

An alternative conceptualization of the model is
Level-1:

$$ATT_{ij} = \beta_{0j} + \beta_{1j} EXP_{ij} + \varepsilon_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_0 + \gamma_1 EXP_{11j} + u_{0j}$$

and

$$\beta_{1j} = \gamma_2.$$

The combined model is

$$ATT_{ij} = \gamma_0 + \gamma_1 EXP_{11j} + \gamma_2 EXP_{ij} + u_{0j} + \varepsilon_{ij}.$$

It follows then that $\gamma_2 = \gamma_L$ is the longitudinal effect, and $\gamma_1 = \gamma_X - \gamma_L$ is the difference between the cross sectional and longitudinal effects. The two models are statistically equivalent. The alternative model clarifies the nature of the assumption in regard to β_{0j}. The model assumes that β_{0j} is uncorrelated with *EXP* at ages 12 to 15. By contrast the models in Equations 15.21 and 15.22 do not make this assumption. The results of the new model are $\hat{\gamma}_2 = \hat{\gamma}_L = .504$ and $\hat{\gamma}_1 = \hat{\gamma}_X - \hat{\gamma}_L = -.095$.

In models considered so far, for both between-subjects and repeated measures data, centering was used to address assumption violations but centering also impacted interpretation of some of the parameters of the model. When the distribution of the level-1 variable is the same for all participants, centering is used to enhance interpretation of the parameters, not to address assumption violations. (The discussion that follows also applies when randomly missing data results in a distribution of the level-1 variable that is not the same for all participants.) For example, in the attitude toward deviant behavior example, attitudes were assessed annually from age 11 to 15 and a level-1 model of interest could be

$$Att_{ij} = \beta_{0j} + \beta_{1j} Age_{ij} + \varepsilon_{ij}, \quad (15.24)$$

where Age_{ij} ranges from 11 to 15. Thus the distribution of age is the same for all participants. The level-2 models are

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

and

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

At any time point, Age_{ij} is the same for all participants and therefore it is impossible for u_{0j} or u_{1j} to be related to age and, consequently, no need to be concerned that u_{0j} or u_{1j} are related to age. Nevertheless, centering can enhance the interpretation of β_{0j} and therefore of γ_{00}. In Equation 15.24 the intercept β_{0j} is the expected attitude score for person j at age 0, and therefore is not subject to a meaningful interpretation. To make the intercept subject to a meaningful interpretation one of two alternative models, obtained by centering *Age*, might be used. The first model is

$$Att_{ij} = \beta_{0j} + \beta_{1j}(Age_{ij} - 11) + \varepsilon_{ij}. \quad (15.25)$$

Using Age_{ij}–11 as the independent variable centers the data so that zero represents the earliest age at which participants are measured. Then the intercept β_{0j} is the expected attitude score for person j at Age_{ij}–11 = 0, that is, at age 11 and in the combined model

$$Att_{ij} = \gamma_{00} + \gamma_{01}(Age_{ij} - 11) + u_{0j} + u_{1j}(Age_{ij} - 11) + \varepsilon_{ij}.$$

γ_{00} is the average attitude at age 11. The second model is

$$Att_{ij} = \beta_{0j} + \beta_{1j}(Age_{ij} - 13) + \varepsilon_{ij}, \quad (15.26)$$

where 13 is the midpoint of the age distribution. Using (Age_{ij}–13) as the independent variable centers the data so that zero represents the midpoint of the ages at which participants are measured. Now, the intercept β_{0j} is the expected attitude score for person j at Age_{ij}–13 = 0, that is, at age 13 and in the combined model

$$Att_{ij} = \gamma_{00} + \gamma_{01}(Age_{ij} - 13) + u_{0j} + u_{1j}(Age_{ij} - 13) + \varepsilon_{ij},$$

γ_{00} is the average attitude at age 13. The intercepts in Equation 15.25 and Equation 15.26 are different and therefore the variance of u_{0j} is different in these models as are the covariances of u_{0j} and u_{1j}. However, γ_{10} is equal in the two models as is the variance of u_{1j}.

When the distribution of the level-1 variable is the same for all participants (or differs only due to randomly missing data), it is common to investigate polynomial trends in the data. For example, a second degree trend in the data could be investigated by replacing Equation 15.24 by

$$Att_{ij} = \beta_{0j} + \beta_{1j}Age_{ij} + \beta_{2j}Age_{ij}^2 + \varepsilon_{ij}. \quad (15.27)$$

Unfortunately, the meaning of β_{1j} is not the same in Equations 15.24 and 15.27. Similarly the meaning of $\gamma_{10} = \mathcal{E}(\beta_{1j})$ is not the same in Equations 15.24 and 15.27. In Equation 15.24, β_{1j} and γ_{10} are the linear trend for person j and the average linear trend in the data, respectively. In Equation 15.27 β_{1j} is the instantaneous rate of change in attitude at zero years of age and γ_{10} measures the average, over adolescents, instantaneous rate of change in attitude at zero years of age. In order to avoid changes in the meaning of terms as higher order powers are added to the model, orthogonal polynomial variables can be used in place of powers of *Age* variable. Table 15.6 shows the orthogonal polynomial variables for use in the *Age* example. For example if the goal were to investigate linear and quadratic trends in the data, the model would be

$$Att_{ij} = \beta_{0j} + \beta_{1j}Linear_{ij} + \beta_{2j}Quadratic_{ij} + \varepsilon_{ij}.$$

TABLE 15.6

Orthogonal Polynomial Variables for the *Age* Example

	Orthogonal Polynomial Variables			
Age	**Linear**	**Quadratic**	**Cubic**	**Quartic**
11	–2	2	–1	1
12	–1	–1	2	–4
13	0	–2	0	6
14	1	–1	–2	–4
15	2	2	1	1

If the data analysts wanted to add a cubic trend to the model, s/he would use

$$Att_{ij} = \beta_{0j} + \beta_{1j} Linear_{ij} + \beta_{2j} Quadratic_{ij} + \beta_{3j} Cubic_{ij} + \varepsilon_{ij}.$$

In both models, β_{1j} measures the linear trend for person *j* and β_{2j} measures the quadratic trend for person *j*. It should be noted that the orthogonal polynomials in Table 15.6 are only appropriate if the values of *Age* are equally spaced or if there are five values for *Age*.

15.16 CONCLUSIONS

In this chapter we have used examples to present the basic results on centering in two-level models. Our approach was to present centering as a method that not only addresses assumption violations but also affects interpretation of parameters. Our presentation was primarily in the context of between-subjects designs. Rather than repeating the developments that apply to both between-subjects designs and repeated measures designs, for repeated measures designs we introduce some issues that apply primarily or uniquely to these designs.

REFERENCES

Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS®*. Cary, NC: SAS Institute Inc.

Baltagi, B. H. (2005). *Econometric analysis of panel data* (3rd ed.). West Sussex, England: Wiley.

Chamberlin, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, *18*, 5–46.

Chamberlin, G. (1984). Panel data. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics* (pp. 1247–1318). Amsterdam: North Holland.

Diggle, P., Heagerty, P., Liang, K-Y., & Zeger, S. (2002) Analysis of Longitudinal Data (2nd ed.). New York: Oxford University Press, USA.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138.

Frees, E. W. (2004). *Longitudinal and panel data*. New York, NY: Cambridge University Press.

Greene, W. H. (2007). *Econometric analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 59*, 1251–1271.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, *24*, 623–641.

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

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

Misangi, V. F., LePine, J. A., Algina, J., & Goeddeke, F. (2006). The adequacy of repeated measures regression for multilevel research designs. *Organizational Research Methods*, *9*, 5–26.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. New York, NY: Holt, Rinehart, and Winston.

Raudenbush, S. W. (1989). "Centering" predictors in multilevel analysis: Choices and consequences. *Multilevel Modeling Newsletter, 1*(2), 10–12.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.).Thousand Oaks, CA: Sage.

Snijders, T. A. B., & Berkhof, J. (2008). Diagnostic checks for multilevel models. In J. de Leeuw and E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 141–175). New York, NY: Springer.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Susser, M. (1994). The logic in ecological: The logic of analysis. *American Journal of Public Health, 85,* 825–829.

APPENDIX A

Mplus Program for Estimating the Random Intercepts Model With Correlation Between Exposure and the Random Intercepts

```
title:
Random intercepts model; Residual
correlated with exposure;
data:
file is "g:\7474\longcross.correct\
cross.dat";
Variable:
names are a1-a5 e1-e5;
MODEL:
   u0i by a1@1;
   u0i by a2@1;
   u0i by a3@1;
   u0i by a4@1;
   u0i by a5@1;
   a1 ON e1 (1) u0i;
   a2 ON e2 (1) u0i;
   a3 ON e3 (1) u0i;
   a4 ON e4 (1) u0i;
   a5 ON e5 (1) u0i;
   [u0i@0];
   u0i with e1;
   u0i with e2;
   u0i with e3;
   u0i with e4;
   u0i with e5;
   [a1 a2 a3 a4 a5] (2);
   a1 a2 a3 a4 a5 (3);
output:
sampstat;
res;
```

APPENDIX B

Screen Shot of HLM Program for Estimating the Random Intercepts Model With Correlation Between Exposure and the Random Intercepts.