­Advanced Topics in Research Methods and Design

Yi Chen

Columbia University

Exercise Three: Hierarchical Linear Models

Note: The SPSS modeling exercises are finished by myself first, then checked and discussed with Jingru Zhang and Lesheng Xu. The proposal is finished completely by myself.

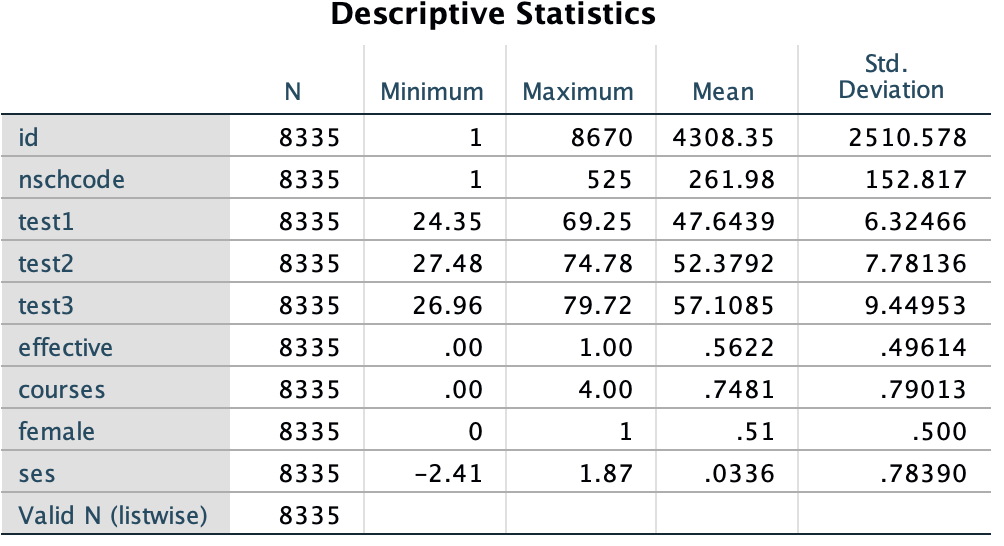
Advanced Topics in Research Methods and Design

# Problem 1: Exercises in Chapter 2

Note: there are too many screen shots I need to make for all plots on the book. In this document, I just screen shot the tables in book since I have to follow the figures on the book to get these tables. I will mention the key process broadly. And some figures in book will be mentioned separately if they are not strongly correlated to the table result. I can provide the output files as well if you want to know the details.

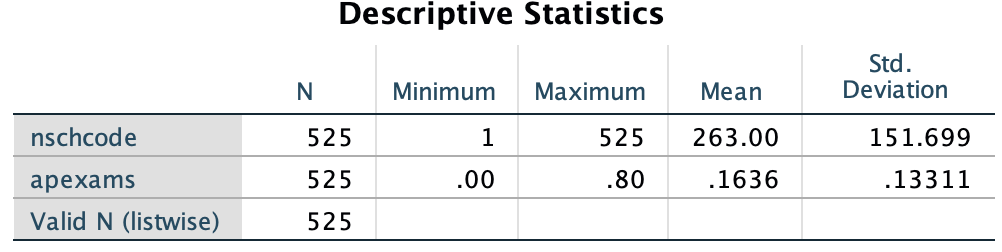
## Table 2.1 and Table 2.2 (Two table are the same)

Notes:

* Data: ch2multivarML1.sav
* Process: analysis 🡺 descriptive statistics 🡺 Descriptives

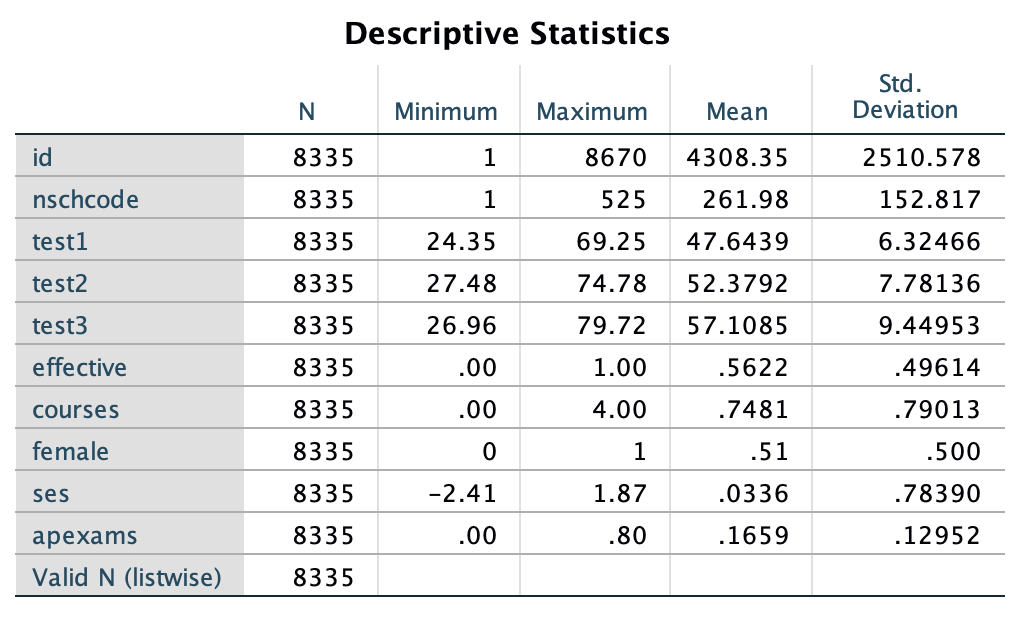
## Table 2.3

Notes:

* Data: apexams.sav
* Process: analysis 🡺 descriptive statistics 🡺 Descriptives

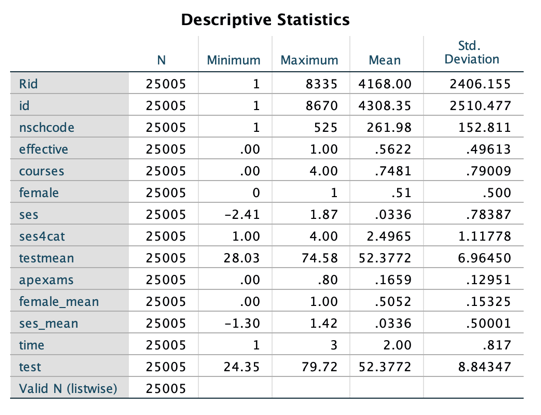
## Table 2.4

Notes:

* Data: ch2multivarML1.sav and apexams.sav
* Process:
  + Sort cases in apexams.sav and ch2multivarML1.sav by nschode
  + Add variable apexams from apexams.sav into ch2multivarML1.sav (merge files by key variable in one to many way)
  + analysis 🡺 descriptive statistics 🡺 Descriptives

## Table 2.5

Notes:

* Data: ch2multivarML1.sav
* Process:
  + Adding new variable by recording (ses4cat)
  + Adding new variable by computing variables (testmean)
  + restructure data structure (from wide format into long format) through making test 1 test2 and test3 into test
  + Descriptive statistics (same as previous tables)

## Figures about compute and rank to recode the level 1 and level 2 data for nested models

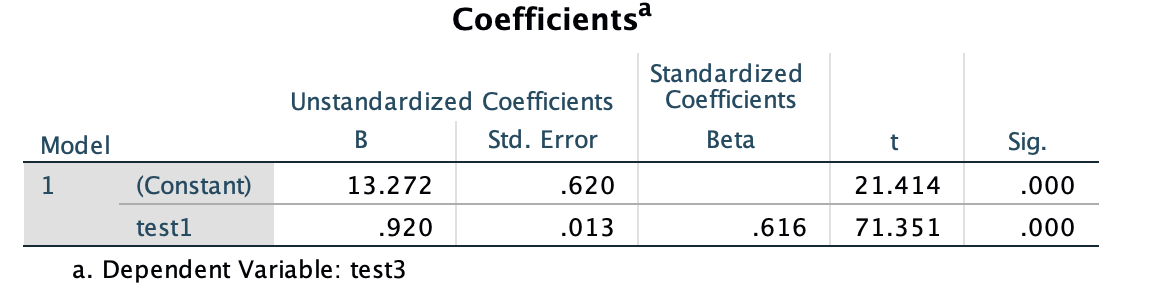
Notes:

* Data: ch2multivarML1.sav
* Process:
  + Clear id and Rid variable
  + Recreate id variable by computer variable using casenum
  + Rank cases of id by nschocode and time togther in sequential racks to unique variables

## Table 2.6

Notes:

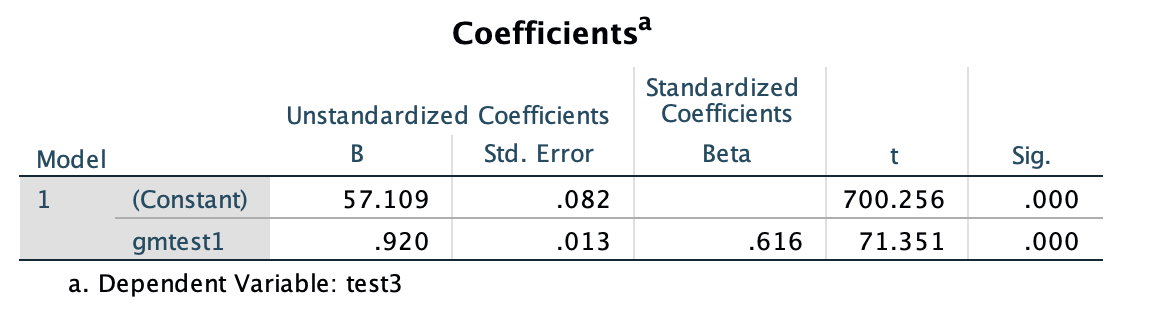
* Data: ch2multivarML1.sav (a new data set, because there is no test 1 and test 3 anymore after data manipulation stated above)
* Process:
  + Regression test 3 on test 1

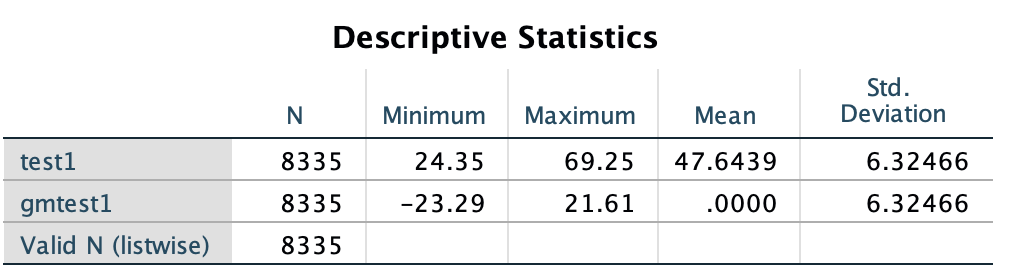


## Table 2.7 and Table 2.8

Notes:

* Data: ch2multivarML1.sav
* Process:
  + Crate new variable gmtest1 (grand mean) by centering the test 1
  + Regression test 3 on gmtest1
  + Descriptive statistics

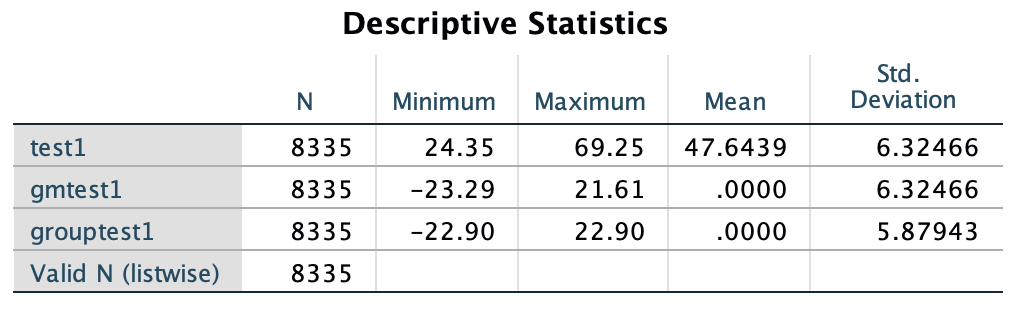




## Table 2.9

Notes:

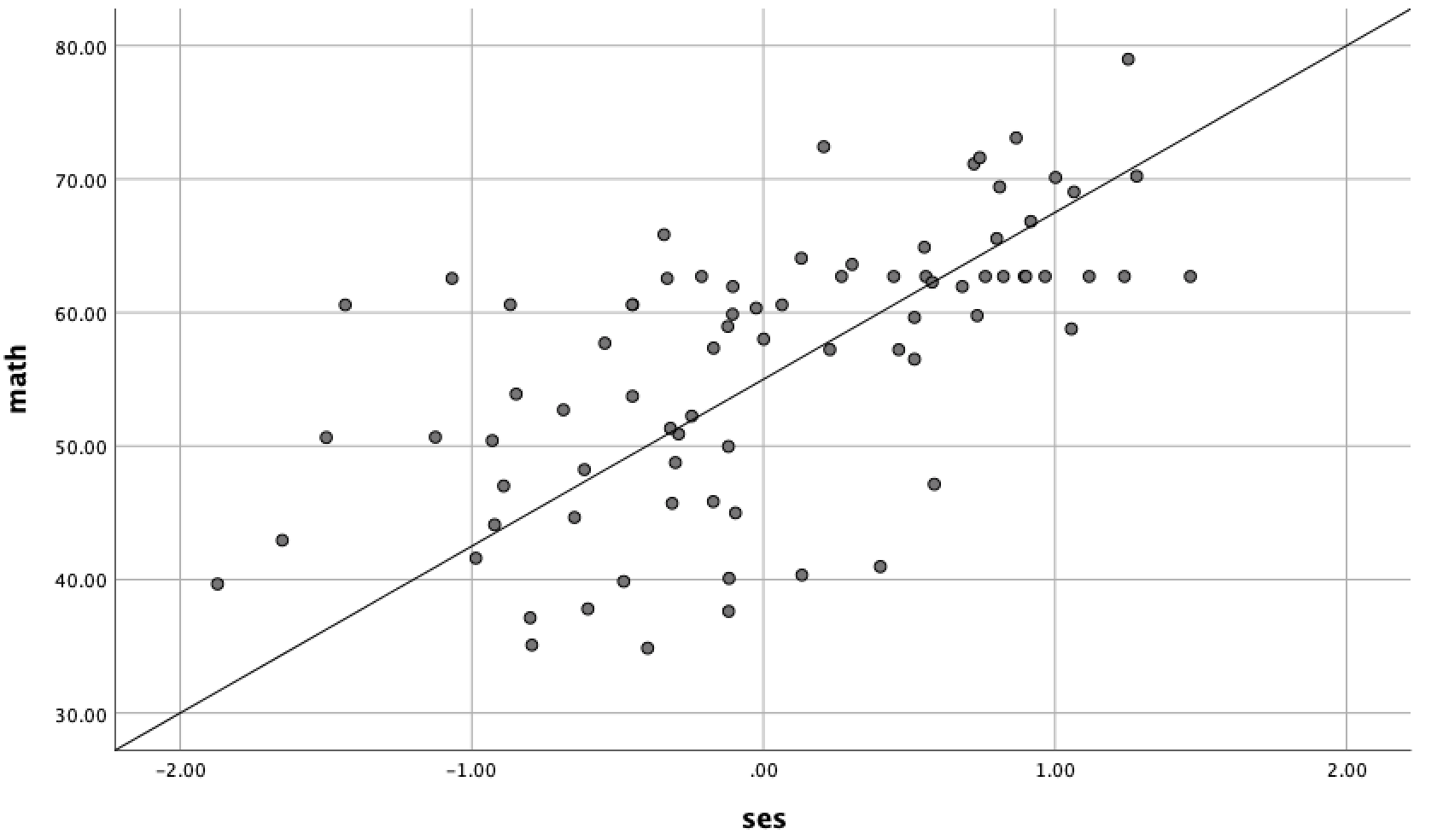
* Data: ch2multivarML1.sav (a new data set, because there is no test 1 and test 3 anymore after data manipulation stated above)
* Process:
  + Aggregate test1 mean by nschocode to add new variable test1gpm
  + Create new variable grouptest1 by test1 - test1gpm
  + Descriptive statistics



# Problem 2: Exercise in Chapter 3

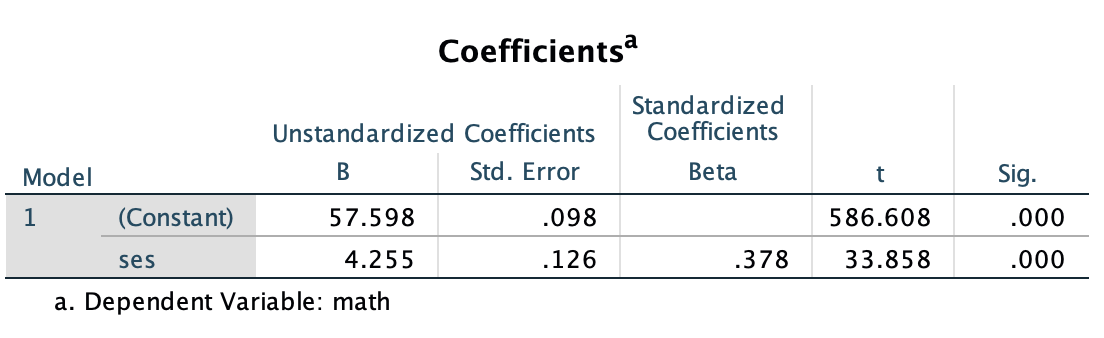
## Figure 3.1

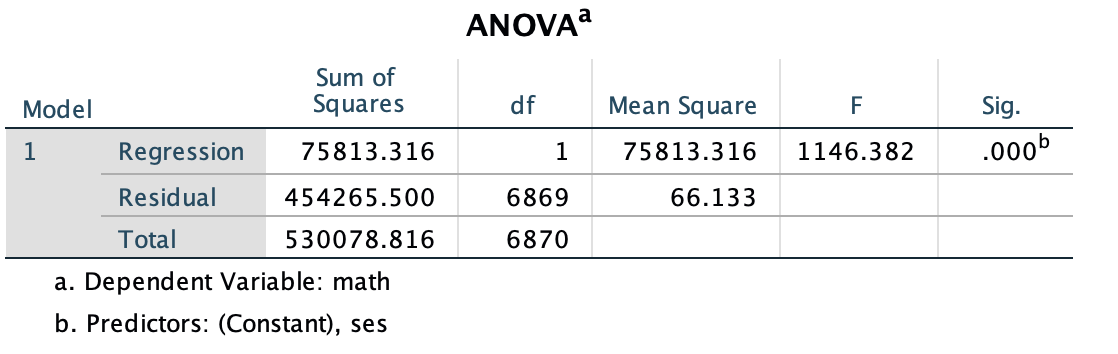
* Data: ch3multilevel.sav
* Process:
  + Select 1-80 cases in the data
  + Scatter plot with regression line



## Table 3.2 and 3.3

* Data: ch3multilevel.sav
* Process:
  + Select all cases
  + Regress math on ses





* Interpretation:

The intercept is estimated as 57.598. This can be interpreted as the sample mean adjusted for individual SES (note that, while not shown, the sample mean unadjusted for SES is 57.734). s

*Y hat* = 57.798 + 4.255 \* *SES，*where *Y* hat indicates that the *predicted* value of *Y* is equal to the estimated intercept plus the coefficient for SES (plus some unknown error). The unstandardized regression coefficient describing the effect of individual SES (beta1) on math achievement is 4.255. The slope coefficient suggests that, on average, as student SES goes up by 1 unit (in this example, 1 *SD*), the predicted student test score would be expected to increase by 4.255 points to 62.053 [57.798 + 4.255(1)]. This is arrived at by multiplying the regression weight (4.255) by the desired 1-unit increase in SES.

We note in the output that in the single-level analysis we also have a standardized regression weight (0.378), which provides a common metric for examining the effects of predictors on the outcome. For a continuous predictor, such as SES in this case, this can be interpreted as a 1-*SD* increase in the predictor producing a 0.378-*SD* increase in the math outcome.

From the table, we can calculate *R*2 as the ratio of the regression variance to the total variance (75813.3/530078.8 = 0.143).

## Table 3.2 and 3.3

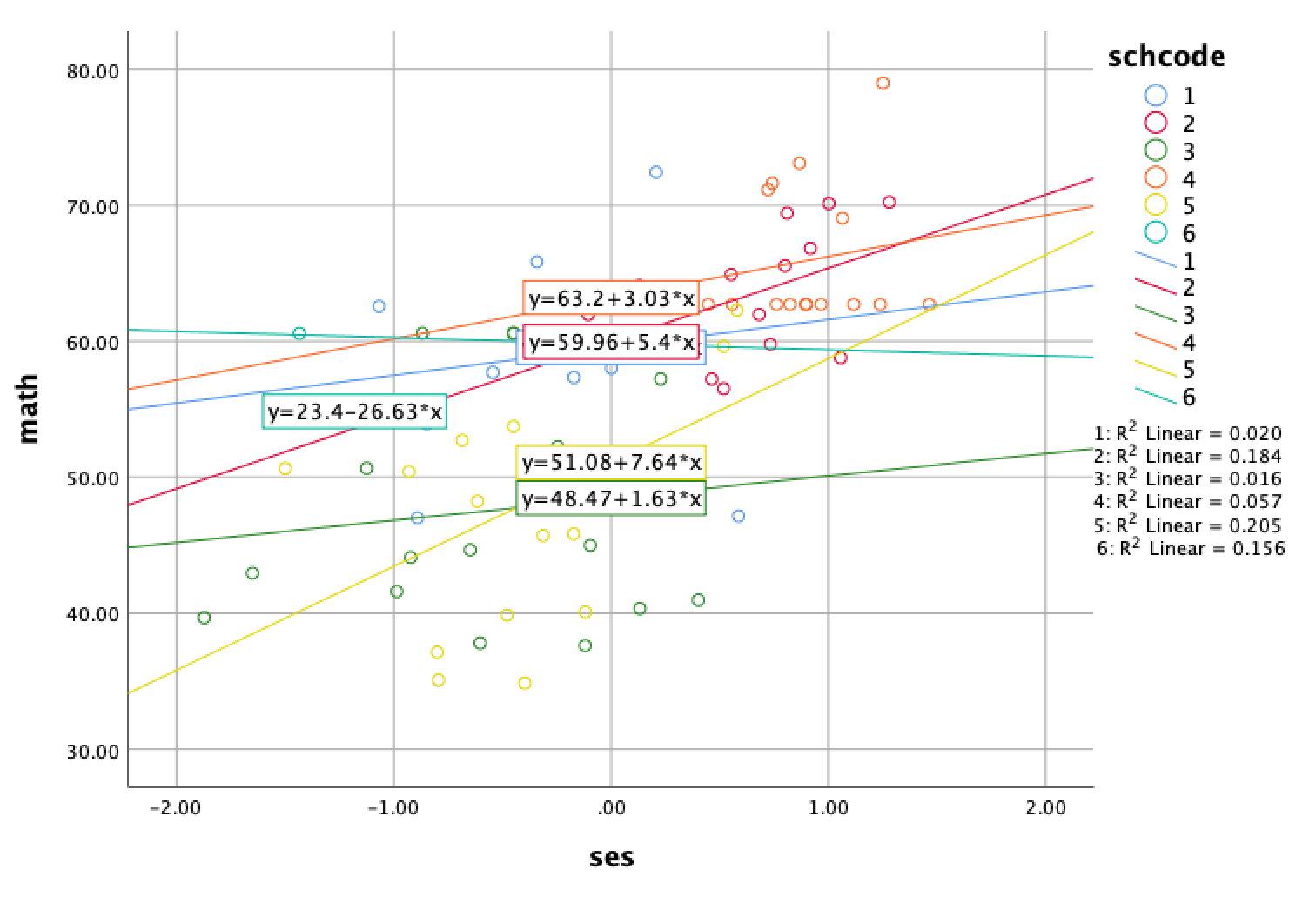
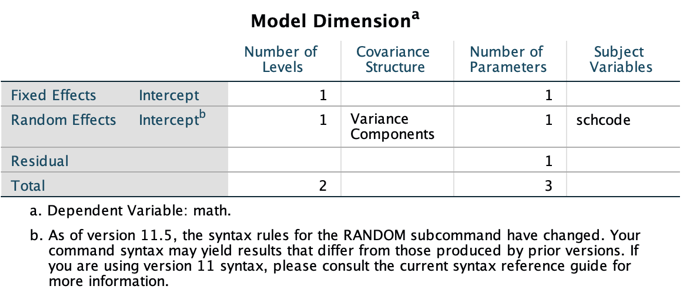
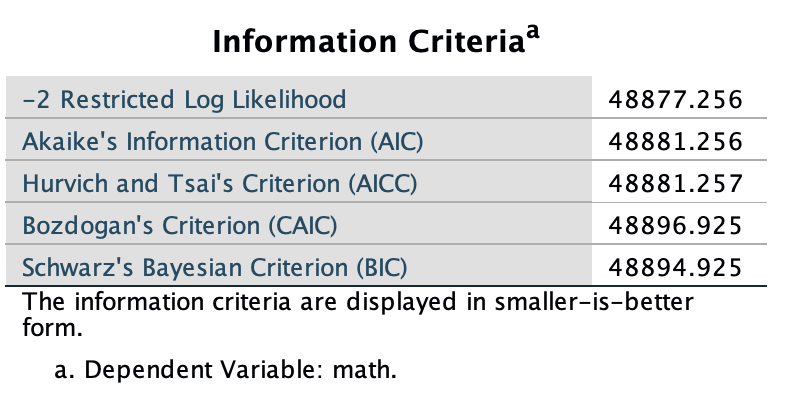
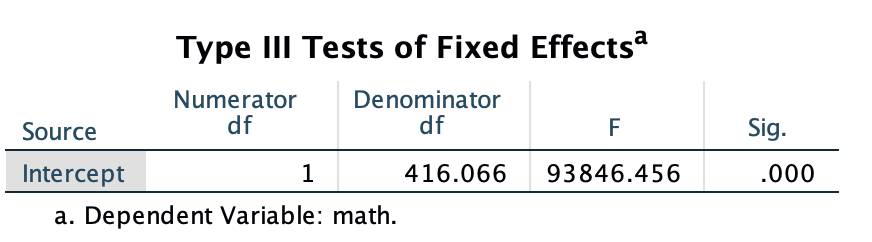
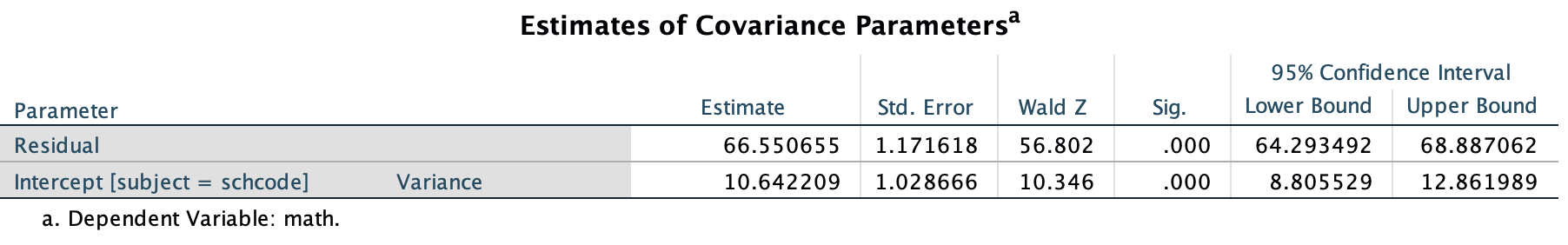
* Data: ch3multilevel.sav
* Process:
  + Select the first 80 cases
  + Scatter plot of math and ses maker by schcode
  + Add regression line makered by schcode
* Interpretation:

Figure above presents the relationship between SES and math achievement for the previous subset of 80 individuals nested in six schools. The figure suggests that the slope relationship accounts for differing amounts of variance within each unit (i.e., with *R*2 coefficients ranging from 0.02 to 0.205). This suggests considerable social distribution of math learning within these six schools.

## Table 3.4 and 3.5

* Data: ch3multilevel.sav
* Process:
  + Select all cases
  + Regression on model 1 (Null)



* Interpretation

The total parameters estimated include the fixed-effect value for the intercept, random Level 2 variance, and the Level 1 variance (referred to as “Residual” in the IBM SPSS output).

The column referred to as “Number of Levels” describes the fixed effects (1) and the number of random effects (1). There is one fixed effect to be estimated (the intercept) and one random effect (the randomly varying intercept). The column referred to as “Subject Variables” indicates the number of levels in the analysis (i.e., in this case, *schcode* [the school identifier] implies a two-level analysis). The covariance structure describes the way the covariance matrix of random effects is at the group level. In this case, we use the default (VC), which provides an estimate of the intercept variance. However, in this first example, at Level 2 there is no random slope variance or covariance between the intercept and slope. In this case, the VC covariance matrix will be the same as an identity matrix.

The variance component output indicates that the proportion of variance in achievement that lies between schools is 0.138. This can be calculated from Equation 3.5 [10.642/ (10.642 + 66.551) = 10.642/77.193], or 13.8%. The intraclass correlation provides a sense of the degree to which differences in the outcome *Y* exist between Level 2 units; that is, it helps answer the question of the existence or nonexistence of meaningful differences in outcomes between the Level 2 units. The results of the null or no-predictors model (basically a one-way ANOVA analysis) suggest that the development of a multilevel model is warranted. Because intercepts vary significantly across schools (Wald *Z* = 10.346, *p* 1 .001), and the ICC suggests that about 13.8% of the total variability in math scores lies between schools, we can develop a multilevel model first to explain this variability in intercepts within and between schools.

Residual parameter describes the variance due to differences among individuals within their respective units. As the table suggests, there is significant variance to be explained within groups (Wald *Z* = 56.802, *p* 1 .001). Similarly, the intercept parameter indicates that the intercepts vary significantly across the sample of schools (Wald *Z* = 10.346, *p* 1 .001). The Wald *Z* test provides a *Z* statistic summarizing the ratio of the estimate to its standard error.

## Table 3.8 to 3.12

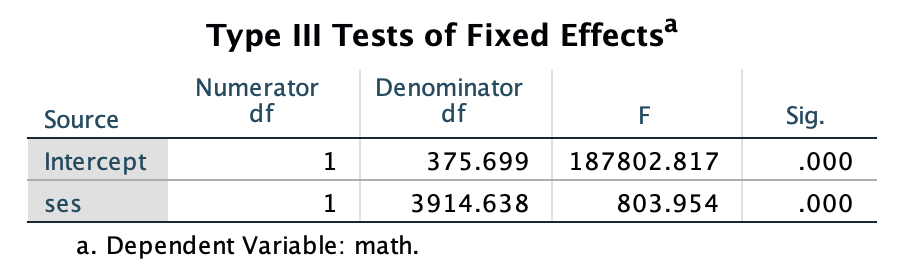
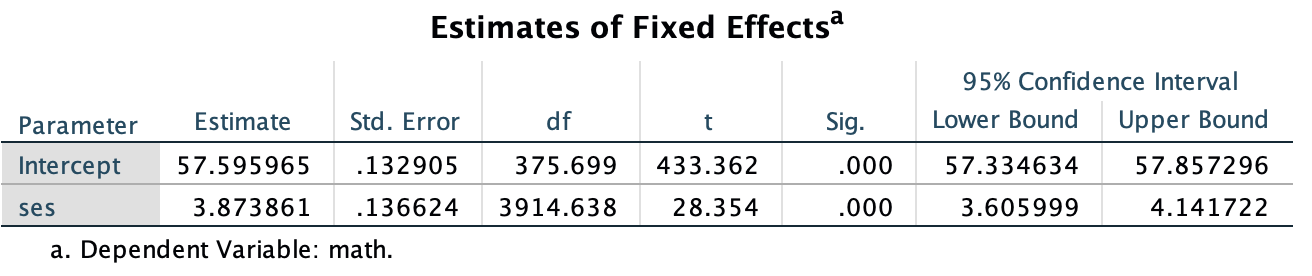
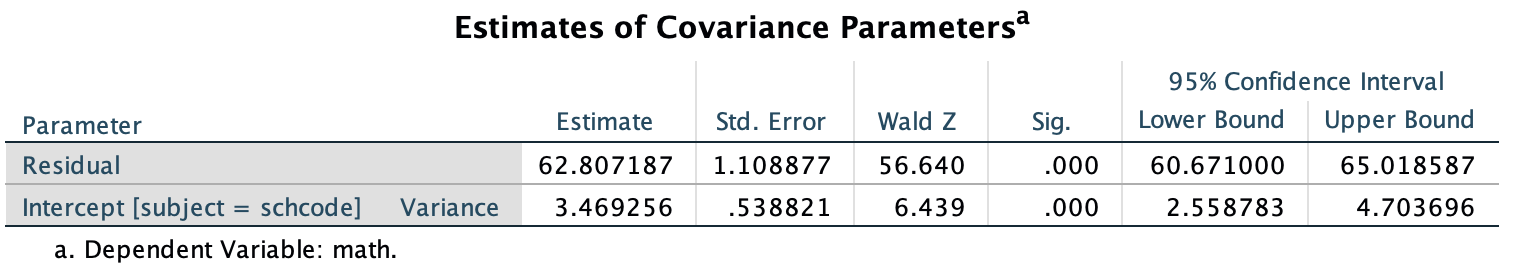
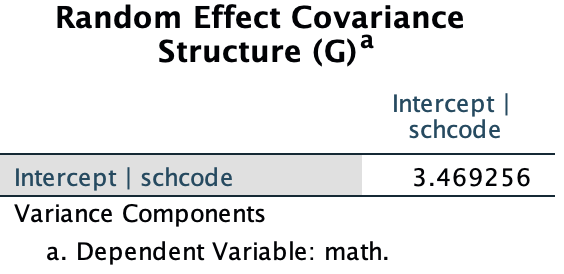
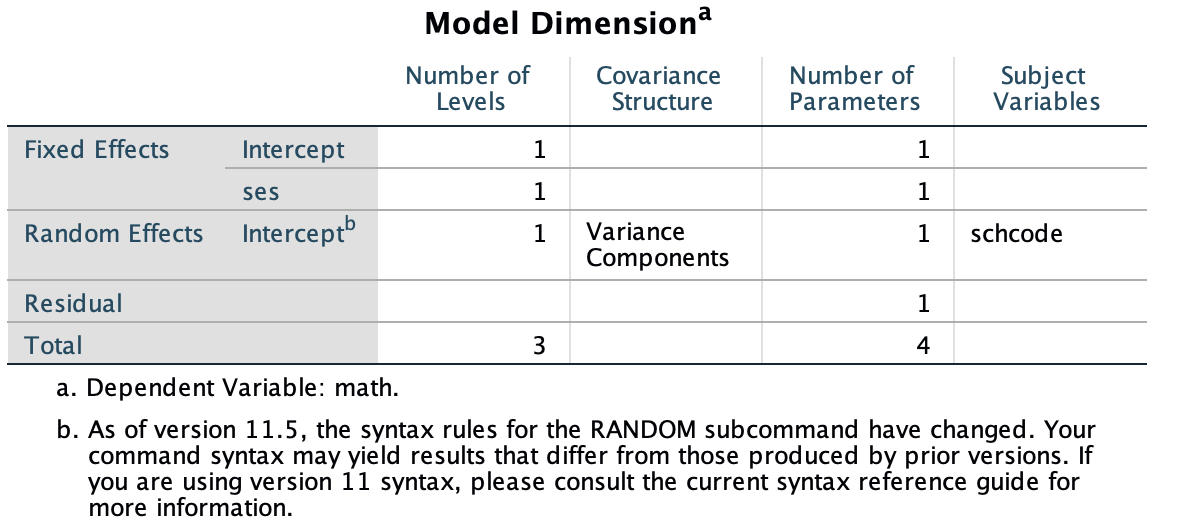
* Data: ch3multilevel.sav
* Process:
  + Select all cases
  + Regression on level one model with covariate ses
* Interpretation

Table 16 summarizes the total number of parameters being estimated (four). This fits with Equation 3.11, suggesting two fixed-effect parameters to be estimated (i.e., the intercept and the within-school predictor SES) and two variance parameters (the random Level 2 variance and the Level 1 residual variance). The column referred to as “Number of Levels” in Table 3.8 describes the fixed effects (two) and the number of random effects (one). There are two fixed effects to be estimated (the intercept and SES) and one random effect (the Level 2 variance component describing variability in the intercept across schools in the sample). The covariance structure describes the way the covariance matrix of random effects is at the group level. In this case, we use the default (VC), which provides an estimate of the intercept variance but again no slope variance because we opted to fix the slope (nor will there be covariance between intercept and slope). This is the same as specifying an identity covariance matrix.

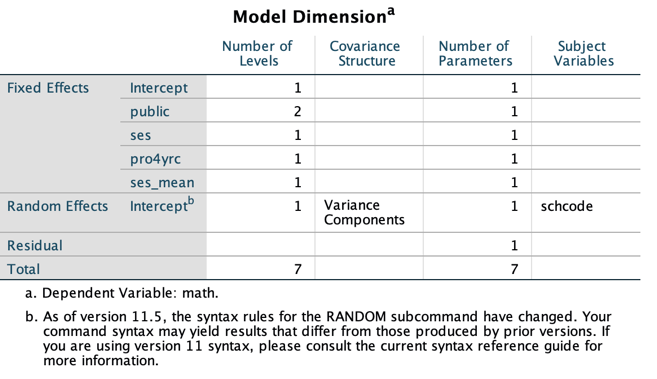
The large *F*-ratio associated with SES in the table suggests that student SES is significantly related to student math scores. The test of the significance of the intercept is generally not of interest, as it is merely a test of whether the intercept is 0 in the model. As the table suggests, we can reject the null hypothesis that the intercept is 0 (i.e., we already know from the null model and descriptive analysis that it is approximately 57.6).

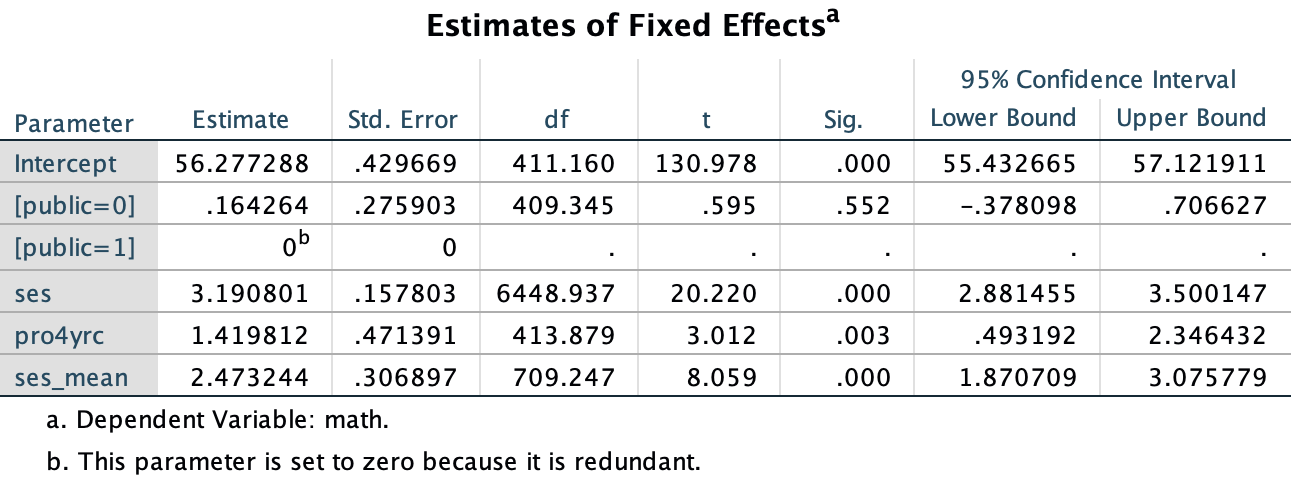
For the result of fixed-effect coefficients, we can see that the intercept (adjusted for student SES) is 57.596. This represents the average school mean adjusted for student SES. The standard error is 0.133. Once again, the *t* test of the significance of the intercept is not really interesting, since it is a test of whether the intercept is equal to 0. In addition, in this case, Table 18 provides the slope for SES (*􏰂* = 3.874) and the standard error (0.137).

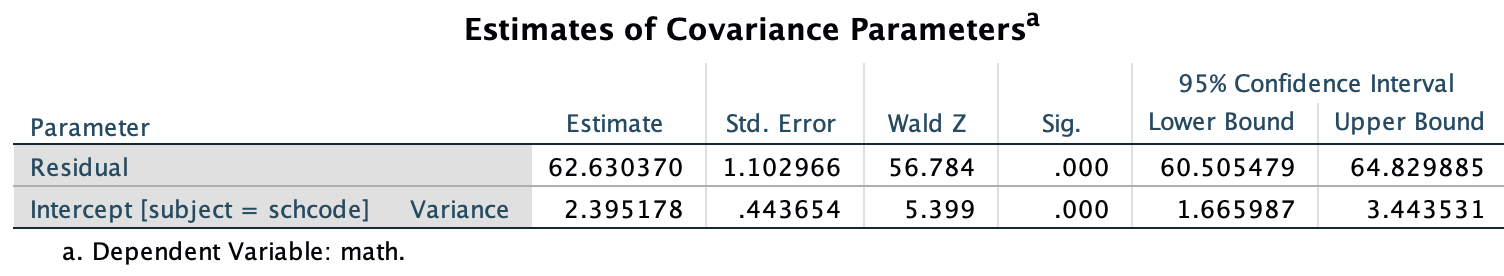
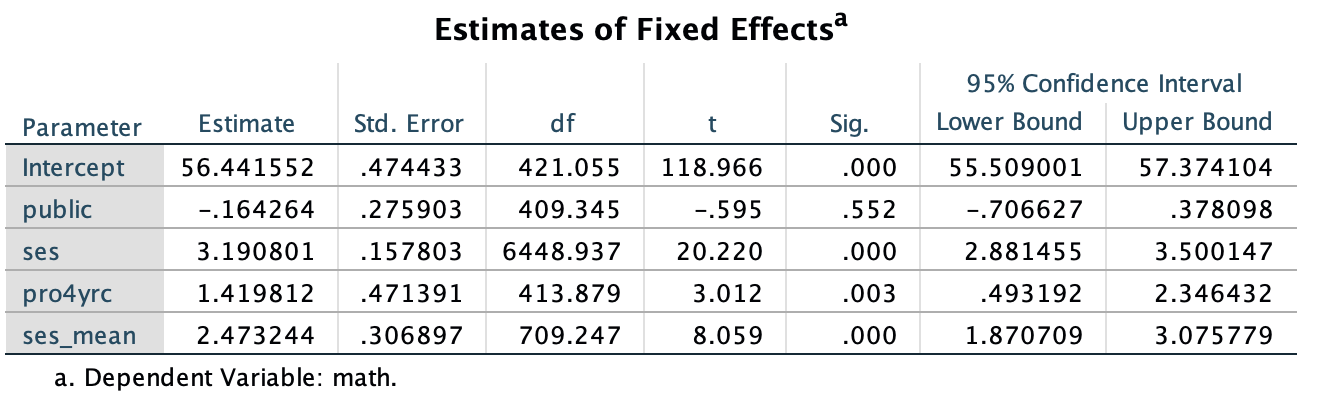
The covariance parameters table also suggests that after the introduction of SES into the model, there is still significant variability to be explained both within schools (Wald *Z* = 56.640, *p* 1 .001) and between schools (Wald *Z* = 6.439, *p* 1 .001). The Wald *Z* test suggests that, even after controlling for student SES within schools, a statistically significant amount of variation in outcomes still remains both within and between schools. This suggests that we could add other predictors (e.g., gender, ethnicity, and motivation) within schools and between schools (e.g., student composition and school process indicators) that might explain this residual variability in intercepts.

The final output provides the random-effect covariance structure for Level 2. In this case, there is only one random effect. Because we specified the SES-achievement slope to be fixed within schools, there is no variance component describing variability in slopes.

## Table 3.12 to 3.16

* Data: ch3multilevel.sav
* Process:
  + Select all cases
  + Regression on group-level





* Interpretation

For the fixed effect table, Level 2 predictive models describe how differences in school variables (e.g., context, resources, student composition, and educational processes) may influence individual processes within each school (Raudenbush & Bryk, 2002), for example, students’ levels achievement outcomes or the relationships between students’ background, motivation, and previous learning and their current outcomes.

We can therefore interpret the intercept as the average school math achievement in *public* high schools (since the reference category in this case is the category coded 1, where the proportion of students planning to attend a 4-year institution is 0, and where school SES composition and student SES are at their respective grand means (0).

Regarding the school-level predictors, controlling for the other predictors in the model, we first note that school type does not affect achievement.

*Y* = 56.28 + 2.47(*ses\_mean* ) + 1.42(*pro4yrc* ) + 0.16(*private* ) + 3.19(*SES* ).

We can interpret the intercept as the adjusted school math score when the predictors in the model are all equal to 0 (i.e., *ses\_mean* = 0, *pro4yrc* = 0, and *student SES* = 0). In the case of *public*, the reference group is actually public schools (coded 1) since when a dichotomous variable is entered in the model as a factor, the reference group is the second category. This suggests that the average school intercept of 56.28 is the predicted score for students in public schools (coded 1) when average school SES (*ses\_mean*) is 0, where no students plan to attend 4-year institutions, and where individual SES is held at the grand-mean average (0).

If we hold the other variables in the predictive model constant, we can estimate the likely school mean if one of the predictors is increased by a unit.

For schools where the proportion of students planning to attend 4-year universities after high school is 0, the expected achievement intercept is 56.28 [56.28 + 1.42(0.0) = 56.28]. In contrast, where all students are planning to attend a 4-year institution the expected math level would be 57.69 [56.28 + 1.42(1.0) = 57.70].

For aggregate SES composition, controlling the other variables in the model, a 1-*SD* increase in aggregate student SES composition above the grand mean would be expected to increase achievement from 56.28 to 58.75 [56.28 + 2.47(1) = 58.75]. We also note in passing that the addition of aggregate student SES background as a school predictor reduces the size of the SES effect on math achievement observed within schools. The model indicates that the effects of student SES background are significant at both the individual and school levels.

We provide MIXED menu commands for treating public school as a covariate (Model 3A) at the end of this section. Readers may also notice that this change results in a slightly different intercept (*􏰄*00 = 56.44 vs. 56.28 in the previous model). This is because the intercept in this cur- rent model now represents the mean for *private* schools (i.e., with public schools [coded 1] now being 0.16 of a point lower). Note that if we add the public-school intercept and private school estimate (56.28 + 0.16), we obtain 56.44, which is the intercept in Table 3.14 when *public* school is entered in the model as a factor instead of a covariate. In Table 3.15, we can see that public schools are 0.16 of point lower than this intercept. This suggests the two model formulations are the same, as we would expect. Since the estimates are the same, we will continue to define school type as a covariate rather than a factor for the rest of the analysis.

The estimates of the variance components suggest that student SES at Level 1 and the three predictors at Level 2 (*public*, *ses\_mean*, and *pro4yrc*) reduce the variance component at the school level substantially (i.e., from 10.64 in the one-way ANOVA model to about 2.40).

# Problem 3: Research Proposal

## Purpose

The general goal for most large-scale assessment in education (e.g. PISA and TIMSS) is to measure the latent proficiency variables (e.g., intelligence, reading skills, math skills, and science skills). However, it may be interesting for the educational researchers and practitioners to identify the valuable indicators, which could best predict the task-takers’ proficiency. Given all the context variables available in the assessment along with their item response, it gives us a chance to make fully use of both item response and background information together through a two-steps model: IRT + Hierarchical linear modeling.

## Literature Review and Theory

Many educational researches suggest that the use of computer (e.g., information searching and social networking) have a positive effect on students’ learning (Brand- Gruwel et al. 2005; Jonassen and Kwon 2001; Tabatabai and Shore 2005). On the other hand, recreational computer use has also been proved by some empirical studies to be beneficial. For example, Bowers & Berland (2013) found that both student use of computers for fun and moderate levels of video gaming were positive and significant on cross-sectional reading and mathematics achievement assessments in high school, controlling for multiple covariates of achievement.

As the importance of computer use becomes more and more obvious, it is a critical question left for the school administrators to decide whether we should invest more money to purchase, update and maintains the computer as one of the school facilities. Bowers & Urick (2011) found no evidence of a direct effect of facility disrepair on student mathematics achievement and instead propose a mediated effects model. Similar to this research, we still lack enough evidence to check whether the availability of computer in school will indeed affect students’ academic performance.

**Research Question**

The requisition questions in this proposal will be: To what extent, the effect of the availability of computers at school affect students’ mathematics proficiency when controlling for general aspects of students’ learning environment.

## Data

The data set I choose is the German sub-sample of the PISA 2012 study (OECD, 2014). The confounding variables at student level includes: students’ gender (), their economic, social and cultural status (), students’ ratings on Classroom management (), and students’ ratings on student-teachers relationship (). The confounding variables at school level is school size (), number of students () and number of computer ().

The data set include a total of 5,001 students nested with in 230 schools. Each school, on average, has 3 to 5 participants. Large-scale assessment usually contains certain number of missing data (both for planned missing by rotation design or the unplanned missing), it may need to use the multiple imputation methods to generate plausible values.

## Method

The two-steps modeling involves IRT model for generate the estimation of students’ mathematics proficiency based on students’ item response in the cognitive testing part. The second step is to take the estimated mathematic proficiency as the dependent variable and all the other variables in interest as the independent variables in hierarchical structure.

In the first step of IRT model (Rash model), the probability of task-taker correctly answer the item can be expressed as:

, where is the estimated latent proficiency and is the estimated latent item difficulty. Based on the model assumption and empirical research results is a continuous variable follows normal distribution. The estimation of latent proficiency is typically based on the marginalized likelihood using EM algorithm.

For the second step of hierarchical linear model part, the latent proficiency can be modeled by all the confounding variables:

, where is the latent proficiency of the task-taker nested in the school , is the grand mean, , and are the coefficient at student level, , and are the fixed effect of school level variables, is the student level residuals, and is the school level residuals.

## Implication

There are two main benefits of this approach: (1) the IRT model provide a more comparable and robust scale of students’ proficiency than directly use the observed achievement data (e.g., total score); and (2) the hierarchical linear structure capture the group level effect (in particular, the most important variable the number of computer () is defined at level 2 in this study), and correlate the violation of residual independence.