How Well do EPGY Curriculum Variables

Predict Student Performance on

the California Standards Test?

Xuejun Shen, Edward Haertel and Patrick Suppes

April 08, 2005

1 Introduction

Pursuant to the No Child Left Behind Act (Pub. Law 107-110), public schools are held accountable for state standardized test scores against prescribed academic standards. In California, student mastery of state standards is assessed using the California Standards Tests (CSTs). Since the Education Program for Gifted Youth (EPGY) curriculum is mapped to California's academic achievement standards, there should be a high correlation between EPGY curricula-based assessments and CST results. The purpose of this study is to explore the degree to which EPGY curriculum-based variables predicted student

CST performances in mathematics. Given the predominantly Title I student sample, this study indicates how useful EPGY curriculum-based assessment variables were for predicting CST scores of disadvantaged students, and thus identifying those who are most likely to need instructional intervention to be successful on the state standards test.

2 EPGY Curriculum-based Variables

The *concept* is the basic unit for instruction and assessment in the EPGY curriculum. The computer-based on-line selection and ordering of concepts and of exercises and items pertaining to each concept, for each student, is a key component of EPGY curriculum-based assessment. This key feature makes it a challenge yet opportunity to explore innovative indicators for student performances on EPGY assessments.

For each concept, a q value is calculated, which represents the probability of an error on that concept. The q value is updated after each item response. When the probability of error--q--is lower than the mastery criterion, .23, the student is judged ready to proceed to items at the next concept level. The final q value when a student finishes with a concept is an important indicator of student performance on the concept, and is included as a predictor variable in the regression analysis. However, the range and variability of q values tend to be limited since the same items (or items in the same concept) are repeatedly presented until students reach the mastery criterion.

Concepts each have a code representing strand, grade level, and placement within that grade. There are five strands of EPGY mathematics curriculum--Algorithms, Fractions, Geometry, Logic, and Measurement--abbreviated as A, F, G, L and M. The final concept that the student reaches in each strand shows the student grade placement within that strand. The average of each student's grade placement levels across these five strands is also included as a variable to predict student performance on the CST.

Another two EPGY curriculum indicators are the average number of items the student answers in each concept, and the average number of concepts the student studies in each strand. The concepts and the exercise items for the concepts vary among students according to their initial grade placement and exercise performances, especially on the initial items. Generally speaking, the better students perform, the fewer items per class they are likely to encounter.

In addition, latency--student response time in seconds--was collected for each item answered. For technical reasons, raw latencies were transformed by taking the natural logarithm of one plus the latency. The transformed latency, averaged over items, is also extracted as a curriculum variable for analysis.

Table 1 briefly summarizes the student-level EPGY variables included in the regression analyses. The mean values of these variables are first calculated by strand. The mean values for the five strands are then averaged, weighted by proportions of concepts among the strands for each student.

Table 1 Summary of Student-Level EPGY Curriculum Variables

Variable Label	Brief Explanation
Mean Q / Student	Mean concept q value for each student across the strands
Mean Ln(1+latency) / Student	Mean ln(1+latency) for each student across the strands
Mean Nr. Of Items / Concept / Student	Mean number of items per concept each student encountered across the strands
Mean Nr. Of Concepts / Strand / Student	Mean number of concepts per strand each student encountered across the strands
Mean Final Grade Placement / Student	Mean final concept in numerical code that the student reached in each strand

3 Sample

The EPGY curriculum was used by students in a middle school in California during Academic Year 2003-2004. The school was classified as a Title I school, and about 1,150 students were enrolled in Grades 6-8 for 2003-2004. The sample of this study was the group of students who participated in the EPGY curriculum. All but one of the sample were Title I students.

As required by the State, beginning in 2004, the students in Grade 8 did not take the same CST math test. Instead, the eighth graders took either the General Math test or a specific-course-based Algebra or Geometry test. Given the relatively small sample size, analyses limited to eighth graders' scores in one of these tests are unlikely to generate valid results. In addition, because high-achieving eighth graders took the Algebra test, the range of ability among eighth graders taking the

Table 2 Frequencies and Percentages of Demographic Variables and Percentages of Demographic Variables by Grade Level

Variable Name	Values	Freq.	Percent. (%)	Percent. (%) for Grade 6	Percent. (%) for Grade 7
Grade in Spring 04	Grade 6	68	42.5		
	Grade 7	92	57.5		
Gender	Male	85	53.1	54.4	52.2
	Female	75	46.9	45.6	47.8
Grade of Entry to District	K	30	18.8	41.2	2.2
	1	7	4.4	10.3	0
	2	2	1.3	2.9	0
	3	1	0.6	1.5	0
	4	12	7.5	17.6	0
	5	6	3.8	7.4	1.1
	6	92	57.5	17.6	87.0
	7	9	5.6	0	9.8
English Fluency	EL (English Learner)	9	5.6	4.4	6.5
	RFEP (Re-designated FEP[2])	24	15.0	14.7	15.2
	IFEP (Initially FEP)	9	5.6	4.4	6.5
	EO (English Only)	117	73.1	76.5	70.7
Ethnicity	Black	73	45.6	44.1	46.7
	Filipino	16	10.0	13.2	7.6
	Hispanic	43	26.9	23.5	29.3
	Other Asian	6	3.8	5.9	2.2
	Other Pacific Isle	1	0.6	0	1.1
	Vietnamese	1	0.6	0	1.1
	White (base ethnic group)	17	10.6	11.8	9.8
Title I Student		159	99.4	100	98.9
GATE[1] Student		12	7.5	10.3	5.4

[1] Gifted and Talented Education

[2] For student English proficiency, English learners are classified into the fluent English proficient (FEP) category based on their performance on The California English Language Development Test.

Table 3 Statistical Summary of Participants' CST Math Scale Scores for Spring 2003 and Spring 2004

CST Scale Scores	Frequency		Mean		Standard Deviation	
Grade in Spring 04	G6	G7	G6	G7	G6	G7
2003 Math	62	89	300.02	289.18	65.93	45.69
2004 Math	66	90	304.91	299.92	49.31	52.38

general math test was restricted. Therefore, only the grade placement test data collected from the 160 students in Grades 6 and 7 were included for analysis. Table 2 presents the numbers of students in each grade level, the overall frequencies and percentages of student demographic variables, and the percentages of these variables by grade level.

Displayed in Table 3 is a summary of 2003 and 2004 CST math scale scores for this group of students by grade. The range of the scale scores for each grade and subject area is 150-600.

4 Regression Analyses

Since the EPGY curriculum and assessment were administered during the 2003-2004 school year, the following analyses show how useful the curriculum variables were for predicting student performances on the Spring 2004 CST in math for Grades 6 and 7 respectively. The analyses employ the regression technique so that the power of each variable in predicting 2004 CST math scores can be evaluated after adjusting for the effects of all the other predicting variables.

The first regression model takes only the EPGY curriculum variables as available predictors, and uses the stepwise procedure to select the significant predictors. That means the only variables included in the model are those for which there would be a less than five percent chance of finding an effect as strong or stronger than the one observed if the true effect were zero, after accounting for all the other variables in the model. Please refer to the appended technical notes for more explanation of the statistical analysis methods and procedures. The second full regression model adds certain student demographic variables, and examines the contribution of each of the student background and EPGY curriculum variables to CST 2004 scores when all these variables are included in the model as predictors. By contrast, the third model adopts the stepwise procedure to choose the most significant predictors from among the full set of both demographic and curriculum variables. Finally, the fourth stepwise model regresses student 2004 CST math scores on EPGY curriculum variables and CST 2003 math scores to identify which EPGY curriculum variables significantly contributed to prediction of the change in student CST math performance from Spring 2003 to Spring 2004.

4.1 Stepwise Regression Model I -- Stepwise Selection of Significant EPGY Curriculum Predictors

As stated above, this stepwise regression model selects the significant curriculum predictors of CST 2004 math scores from among all the extracted student-level curriculum variables for each grade level. Table 4 presents the standardized coefficients for the significant predictors. As shown in the table, for Grade 6, the significant predictors are the number of items per concept, and mean latency. The standardized coefficients for these two variables, -.495 and -.320

respectively, show that for an increase in number of items per concept or response latency by one standard deviation, CST 2004 math scores will decrease by .495 or .320 of a standard deviation. This means that the fewer items the students encountered per concept and the faster the students answered the questions, the higher they achieved on the CST. As for Grade 7, besides response latency and number of items per concept, student final grade placement also emerges as a significant predictor. The model adjusted R-squared values, .522 for Grade 6 and .450 for Grade 7, as shown in the last row of Table 4, indicate the overall proportion of CST 2004 math score variance that can be predicted by the sets of predictors.

Table 4 Standardized Regression Coefficients of EPGY Curriculum Variables Selected by the Stepwise Procedure (Dependent Variable: 2004 CST Math Scale Score)

Available EPGY Curriculum Variables	**Standardized Regression Coefficients of Significant Predictors**	
	Grade 6 (N=55)	**Grade 7 (N=84)**
Mean Q / Student		
Mean Ln(1+latency) / Student	-.320	-.348
Mean Nr. Of Items / Concept / Student	-.495	-.276
Mean Nr. Of Concepts / Strand / Student		
Mean Final Grade Placement / Student		.300
Model Adjusted R-Squared	.522	.450

Table 5 Standardized Regression Coefficients of the Full Set of Demographic and EPGY Curriculum Variables (Dependent Variable: 2004 CST Math Scale Score)

Demographic and EPGY Curriculum Variables	**Standardized Regression Coefficients of Full Set of Predictors**	
	Grade 6 (N=53)	**Grade 7 (N=82)**
Demographic Variables		
Gender	.141	.216*
Grade of Entry to School	-.132	-.100
Is Title I		.041
Is GATE	.081	.162
Is Indian	-.088	
English Fluency	.038	.084
Ethnicity		
Is Black	-.607*	-.401*
Is Filipino	-.118	.034
Is Hispanic	-.275	-.081
Is Other Asian	-.066	-.030
Is Other Pacific Isle		-.142
Is Vietnamese		.025
Curriculum Variables		
Mean Q / Student	-.013	-.030
Mean Ln(1+latency) / Student	-.183	-.243*
Mean Nr. Of Items / Concept / Student	-.416*	-.247
Mean Nr. Of Concepts / Strand / Student	-.158	-.046
Mean Final Grade Placement / Student	.141	.276*
Model Adjusted R-Squared	.639	.571

* Significant predictors with probability of contributing to CST 2004 math scores within .05 significance criterion

4.2 Regression Model II -- Full Set of Demographic and EPGY Curriculum Variables as the Predictors

In this model, in addition to EPGY curriculum variables, the demographic variables are also included as predictors of CST 2004 math scores. The background demographic variables collected for the student sample include gender, ethnicity, grade of entry to the present school (an indicator of student mobility), English proficiency, and whether he or she is a GATE, Title I or Indian student. Table 5 illustrates the standardized regression coefficients of all the predictors for both grade levels. For Grade 6, among the full set of predictors, whether the student is a Black student and the number of exercise items per concept are statistically significant; As for Grade 7, the significant demographic variables are Gender and whether the student is Black, and the significant curriculum variables are the response latency and final grade placement across the strands for each student.

4.3 Regression Model III -- Stepwise Selection of Significant Demographic and EPGY Curriculum Predictors

For this model, the stepwise procedure is applied to the same set of variables included in the preceding full model, to select just the significant variables contributing to CST 2004 math scores. Table 6 displays the standardized regression coefficients of the significant predictors for each grade level. As shown in the table, the significant predictors selected for Grade 6 are whether the student is GATE, whether the student is Black or Hispanic, and the mean number of items per concept class. Those selected for Grade 7 are grade of entry to school, English fluency level, whether the student is Black, response latency, number of items per concept, and final grade placement. The adjusted R-squares for the models are .657 and .577, indicating the models account for 65.7 percent and 57.7 percent of the variation in CST 2004 math scores for Grade 6 and for Grade 7, respectively.

Table 6 Standardized Regression Coefficients of Demographic and EPGY Curriculum Variables Selected by the Stepwise Procedure (Dependent Variable: 2004 CST Math Scale Score)

Available Demographic and EPGY Curriculum Variables	Standardized Regression Coefficients	
	Grade 6 (N=54)	Grade 7 (N=82)
Demographic Variables		
Gender		.204
Grade of Entry to School		
Is Title I		
Is GATE	.199	
Is Indian		
English Fluency		-.276
Ethnicity		
Is Black	-.448	-.276
Is Filipino		
Is Hispanic	-.354	
Is Other Asian		
Is Other Pacific Isle		
Curriculum Variables		
Mean Q / Student		
Mean Ln(1+latency) / Student		-.310
Mean Nr. Of Items / Concept / Student	-.470	-.262
Mean Nr. Of Concepts / Strand / Student		
Mean Final Grade Placement / Student		.202
Model Adjusted R-Squared	.657	.577

In Model I, using only the EPGY curriculum variables, the percent of variance accounted for (R squared) was about .5. Models II and III, which also included demographic variables, brought the R squared up to around .6. In the following model, when the prior year's test scores were admitted as an predictor of 2004 CST scores, along with EPGY curriculum variables, the R squared value of the model rose to a much higher level.

4.4 Regression Model IV -- Stepwise Selection of Significant EPGY Curriculum Variables and 2003 Math Scale Score Predictors

These regression models explore how much variation of CST 2004 math scores can be explained by the selected curriculum variables and the CST scores of the previous year--2003. When student achievement in 2003 was held constant, demographic variables did not appear to be significant predictors of CST 2004 scores except the one for the ethnicity of Filipino for the seventh grade. From the pool of CST 2003 scores and EPGY curriculum variables, the stepwise regression procedure selects, in addition to 2003 scores, response latency as the significant predictor of CST 2004 scores for Grade 6, and response latency and mean final grade placement for Grade 7. The subsequent analysis thus focuses on how much mean final grade placement and response latency can explain the change of student CST scores from 2003 to 2004.

As stated above, the final grade placement variable is not a significant predictor of CST 2004 scores in the linear regression model for Grade 6; however, there seems to be a nonlinear relationship between these two variables as shown in the scatter plot in Appendix D. A quadratic term of the final grade placement is thus added to represent its nonlinear relationship with CST 2004 scores, and appears to be a significant predictor as shown in Table 7, which displays the standardized regression coefficients of the final models fitted for both grade levels.

In addition to the individual effects of the final grade placement and response latency, also explored are the interaction effects of these two variables with CST 2003 math scores, since it is often the case that students with higher prior achievement tend to absorb the curriculum and respond to the assessment at a faster pace. However, after a series of model fitting, we find that none of the interaction terms are significant predictors. As shown in Table 7, for Grade 6, although mean response latency and the interaction effect between response latency and CST 2003 scores are not significant predictors, they respectively raise the adjusted R-square value by .017 and .09. These two variables are thus included in the model for Grade 6, whose overall adjusted R-square value reaches .817. As to Grade 7, the linear regression model without interaction terms seems to fit the data best, and both mean final grade placement and response latency are significant predictors when CST 2004 scores are regressed on them and CST 2003 scores. The model for Grade 7 can explain 84.2% of the variability of CST 2004 scores.

The high predictive power of the models indicates that student mean final grade placement and student response latency can largely explain the change in student CST scores from Spring 2003 to Spring 2004. The following figures show, for each grade level, the scatter plots of observed versus predicted CST 2004 scale scores with 95% prediction interval bands, which should contain 95% of the observed 2004 CST scores. The prediction interval is an indicator of how the observed scores deviate from the fitted regression line. As shown in the graphs, the standard errors of the estimate of CST 2004 scores for both grades are close, 20.544 for Grade 6, and 20.110 for Grade 7.

Table 7 Standardized Regression Coefficients of and 2003 CST Math Scale Score Predictors Selected by the Stepwise Procedure (Dependent Variable: 2004 CST Math Scale Score)

Available EPGY Curriculum and CST 2003 Math Score Variables	Standardized Regression Coefficients	
	Grade 6 (N=55)	**Grade 7 (N=81)**
CST 2003 Math Scale Score	2.063*	.787*
Mean Final Grade Placement / Student	-2.488*	.106*
Squared Mean Final Grade Placement / Student	2.522*	
Mean Ln(1+latency) / Student	.460	-.154*
Interaction Effect Between Mean Ln(1+latency) and CST 2003 Math Scale Score	-1.110	
Model Adjusted R-Squared	.817	.842

* Significant predictors with probability of contributing to CST 2004 math scores within .05 significance criterion

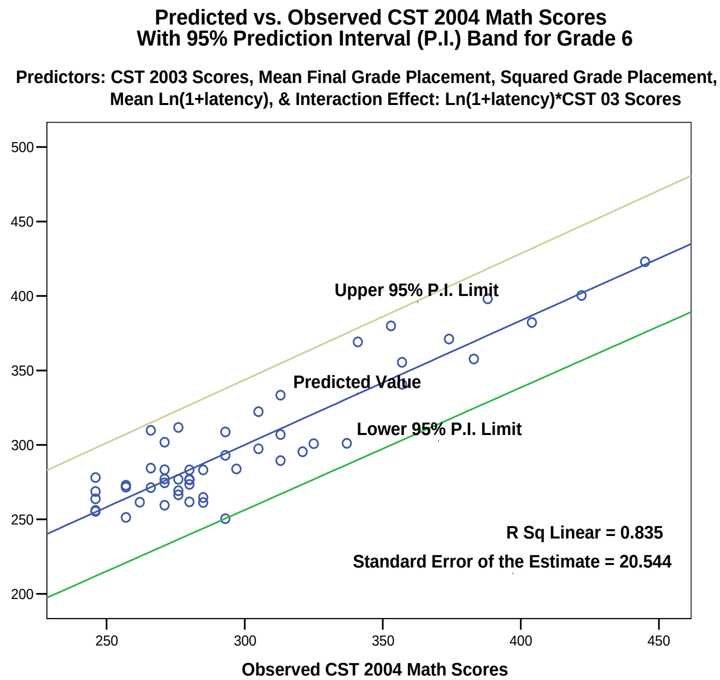


Figure 1. Predicted vs. Observed CST 2004 Math Scores for Grade 6 with 95% Prediction Interval

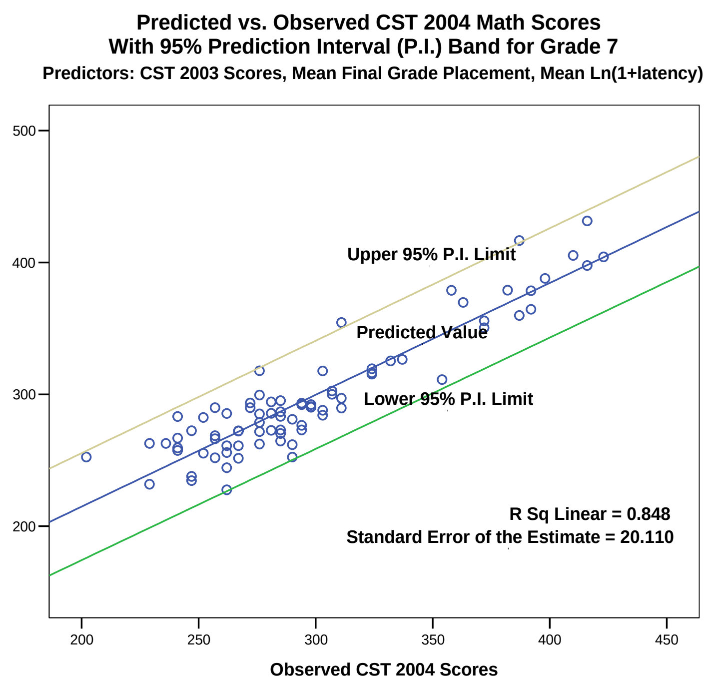


Figure 2. Predicted vs. Observed CST 2004 Math Scores for Grade 7 with 95% Confidence Intervals

4.5 Predicted vs. Observed Number of Students at 2004 CST Math Performance Levels

CST scores are categorized into five performance levels: Far Below Basic, Below Basic, Basic, Proficient, and Advanced, based on designated cut-off scores for 2004 (which are shown in Appendix C). The scores predicted by the preceding regression models, with CST 2003 scores, final grade placement (with its quadratic term for Grade 6) and response latency (with its interaction effect with CST 2003 scores for Grade 6) as predictors, are also assigned to those five performance levels according to the same cut-off scores. Table 8 and Table 9 are the cross-tabs comparing the predicted versus observed number of students below, and at or above the Basic performance level. The predictions seem to be quite accurate, with 90.2% and 93.8% of them consistent with the true observations.

Table 8 Predicted vs. Observed Grade 6 Students at All Performance Levels for CST 2004 Math

Predicted \ Observed	Frequency		Correct Prediction Rate
	Below Basic	Basic or Above	
Below Basic	29	3	92.8%
Basic or Above	4	17	81.0%
		Total	*90.2%*

Table 9 Predicted vs. Observed Grade 7 Students at All Performance Levels for CST 2004 Math

Observed / Predicted	Frequency: Below Basic	Frequency: Basic or Above	Correct Prediction Rate
Below Basic	51	4	85.7%
Basic or Above	1	24	98.1%
		Total	*93.8%*

Table 10 and Table 11 further display the predicted versus observed numbers of students for all performance levels. The correct prediction rates, overall, reach 74.5% and 70% respectively for each grade level. The prediction accuracy rates seem to be affected by the fact that the predictions are more accurate when the scores are closer to the means, since the variance of predicted values is smaller than that of observed values.

Table 10 Predicted vs. Observed Grade 6 Students Below vs. At or Above Basic Performance Level for CST 2004 Math

Observed / Predicted	Frequency: Far Below Basic	Below Basic	Basic	Proficient	Advanced	Correct Prediction Rate
Far Below Basic	3	4	0	0	0	57.1%
Below Basic	2	20	1	0	0	87.0%
Basic	0	4	6	0	0	60.0%
Proficient	0	0	2	7	0	77.8%
Advanced	0	0	0	2	2	50.0%
					Total	*74.5%*

Table 11 Predicted vs. Observed Grade 7 Students Below vs. At or Above Basic Performance Level for CST 2004 Math

Predicted \ Observed	Frequency					Correct Prediction Rate
	Far Below Basic	Below Basic	Basic	Proficient	Advanced	
Far Below Basic	5	6	0	0	0	38.5%
Below Basic	8	32	4	0	0	82.1%
Basic	0	1	8	1	0	61.5%
Proficient	0	0	1	10	2	83.3%
Advanced	0	0	0	1	1	33.3%
					Total	*70.0%*

5 Conclusion

In summary, among the demographic variables included for analysis, whether students are in the GATE program, ethnicity, student gender, and English fluency level emerge as significant predictors of student CST performances at the 6^{th} and 7^{th} grade levels. As to the EPGY curriculum variables, the significant predictors are mean response latency, mean number of items per concept, and final grade placement across the five concepts for each student. Combined with student performance on 2003 CST, student mean final grade placement in the EPGY math curriculum, along with its quadratic term, predominantly predicts student performances on CST 2004 math for Grade 6. As to the final model with CST 2003 scores as a predictor for Grade 7, in addition to final grade placement, mean transformed response latency also seems to be significant in predicting CST 2004 scores. These two final models account for over 80% of the variation of

CST 2004 scores. Perfect prediction is impossible, if only because CST scores themselves have some measurement error. The theoretical limit to the R-squared value would be given by the reliability of CST, which is probably less than .9.

Although response latency has gained attention in the psychological research, it is not a conventional variable collected in educational assessment programs. Latency involves the whole question answering process from the presentation of an exercise to finally reaching an answer. In EPGY curriculum, latency is an indicator of a characteristic individual difference among students, and it has been observed that overall slower students make more errors than their peers in the early grades (Grades One to Three) and fewer errors than their peers in the later grades (Grades Four to Six) (Tock & Suppes, 2002). The theoretical and practical implication of the significant effect of latency, especially the actual factors that affect its length, awaits further exploration.

The significant predictive power of the EPGY curriculum variables, especially the mean final grade placement across the math strands, confirms the positive correlation between student performances on EPGY curriculum-embedded assessments and the CST. Since the CST addresses the state academic standards, the correlation between EPGY and CST assessment results tends to confirm the alignment between the state mathematics standards and EPGY curriculum. Since all but one of the study subjects are Title I students, EPGY curriculum-based assessments seem to be effective for identifying students from lower socioeconomic backgrounds who are likely to need some additional instructional intervention to be successful on the CST. EGPY curriculum assessments should become better predictors of CST scores as students spend more time working through the curriculum. It would be interesting and rewarding to keep monitoring and experimenting with EPGY curriculum variables as predictors of external assessment results, along with the delivery of the EPGY curriculum.

References

Chatterjee, S., Hadi, A. S., & Price, B. (2000). *Regression analysis by example* (3rd ed.). New York: John Wiley & Sons.

Tock, K., & Suppes, P. (2002). *The high dimensionality of students' individual differences in performance in EPGY's K6 computer-based mathematics curriculum.* Retrieved January 12, 2005, from http://www-epgy.stanford.edu/research/trajectories.pdf

Appendix A
Technical Notes

Mean

The sample mean m of a set of *n* observations from a given distribution is defined by

$$\frac{1}{n}\sum_{k=1}^{n} x_k \ .$$

It is an unbiased estimator for the population mean µ. (Quote from Mathworld at http://mathworld.wolfram.com/Mean.html)

Standard Deviation and Variance

The standard deviation is the most commonly used measure of how spread out a distribution is. It is computed as the square root of the variance, which is the average squared deviation of each number from its mean $\bar{x}$:

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}.$$

It is a common practice to subtract one from N when the sample standard deviation is treated as an inferential statistic of a population:

$$s_{N-1} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}.$$

Correlation

Correlation is the linear association between two random variables. It is usually measured by a correlation coefficient, such as Pearson's r, which indicates the proportion of the variability in one variable accounted for by the covariance of the two variables:

$$\mathrm{r} = \frac{1}{n-1}\sum\left(\frac{x-\bar{x}}{s_x}\right)\left(\frac{y-\bar{y}}{s_y}\right).$$

In the formula above, $\bar{x}$ and $\bar{y}$ are the means of the observations of *x*, *y* variables; s_x and s_y stand for their standard deviations; and *n* is number of observations. The absolute value of the coefficient shows the magnitude of correlation, and the sign of the coefficient indicates whether one variable increases in the same direction as the other one or not.

Multiple Linear Regression

Regression analysis uses one or more predictor/explanatory variables to estimate or predict the response/dependent variable. The multiple linear regression model expresses the linear relationship between the response variable and the set of predictors as follows:

$$Y = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p + \varepsilon,$$

where *Y* stands for the response variable, $X_1, \ldots, X_p$ denote the set of predictor variables, $\beta_0, \beta_1, \beta_2, ..., \beta_p$ are the partial regression coefficients (or simply "regression coefficients"), and ε represents a random error, which is assumed to account for the discrepancy between the true/observed value Y and the estimated/predicted value $\hat{Y}$ of the response variable.

Regression Coefficients and R Squared

The regression analysis results largely indicate the variability of the response variable accounted for by the total set of predictors in the model, as well as the relative importance of each individual predictor. The relative importance of predictors is indicated by the partial regression coefficients. The most common method of estimating the coefficients is called the least squares method, which minimizes the sum of squares of the errors or residuals--the difference between the fitted and observed response variable values. The regression coefficients are analogous to correlation coefficients, and indicate the change of the response variable corresponding to the change in a certain predictor. However, the distinct feature of the regression coefficient is that it corrects for the variability of *Y* explained by other predictor variables. To take the regression model with two predictors as an example, the formula for computing the partial coefficients is

$$\beta_1 = \frac{r_{Y1} - r_{Y2} r_{12}}{1 - r_{12}^2} \text{ and } \beta_2 = \frac{r_{Y2} - r_{Y1} r_{12}}{1 - r_{12}^2},$$

where r_{Y1} and r_{Y2} stand for the correlation coefficients between *Y* and each predictor, and r_{12} is the correlation between the two predictors.

β_0 is the intercept of the regression model, and estimates the mean Y value without the influence of the predictor variables:

$$\beta_0 = y - \beta_1 x_1 - \beta_2 x_2 .$$

The total proportion of *Y* variance explained by the model is indicated by R squared, which is the square of the correlation between the estimated and observed values of the response variable:

$$R^2 = [\, Cor(Y, \hat{Y})\,]^2,$$

and

$$Cor(Y, \hat{Y}) = \frac{\sum (y_i - \overline{y})(\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum (y_i - \overline{y})^2 \sum (\hat{y}_i - \overline{\hat{y}})^2}} ,$$

where $\overline{y}$ is the mean of observed response variable values, and $\overline{\hat{y}}$ is the mean of predicted values.

As to adjusted R squared, it takes account of the number of predictor variables included in the regression model:

$$R_a^2 = 1 - \frac{n-1}{n-p-1}(1 - R^2) ,$$

where n is the number of observations, and p is the number of predictor variables.

Prediction/Confidence Interval and Standard Error of Prediction

The prediction intervals are the predicted response variable values plus or minus a certain multiple of the standard error of each prediction. The 95% confidence interval bands are higher or lower than the predicted values by about twice the standard error. The standard error is an indicator of the deviation of the observed response variable values from the predicted or estimated values. As an illustration, following is the formula to compute the standard error of prediction in a regression model with a single predictor:

$$s.e. = \hat{\sigma}\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}}\ ,$$

where x_0 is the mean predictor value, and $\hat{\sigma}$ is an unbiased estimate of the unknown parameter σ:

$$\hat{\sigma} = \sqrt{\frac{\sum e_i^2}{n-p-1}} = \sqrt{\frac{\sum(y_i-\hat{y}_i)^2}{n-p-1}}\ .$$

In the formula above, y_i is the observed response variable value, $\hat{y}_i$ is the estimated value, n is the number of observations, and p is the number of predictors.

Stepwise Variable Selection Method

An important goal in regression analysis is to arrive at adequate description of observed phenomenon in terms of as few meaningful variables as possible (Chatterjee, Hadi & Price, p. 68). The criterion for a predictor to be significant is that its regression coefficient is significantly different from zero.The usual practice is to set the significant probability level at p = .05, which means: Under the null hypothesis that the true regression coefficient is exactly zero, there is less than a 5 percent chance of obtaining an observed coefficient as large or larger in absolute value than the one observed. The *p*, probably value, is obtained by conducting a t-test of the partial regression coefficient for the predictor,

$$t = \frac{\hat{\beta}}{s.e.(\hat{\beta}\ )},$$

where the standard error (s.e.) is an estimate of the standard deviation of the distribution of values of that regression coefficient that would be obtained with repeated samples of the same size from the same population, conditional on the values of all the independent variables. Again, for illustrative purpose, following is the formulae to compute the standard error for the intercept, $\hat{\beta}_0$ and for the regression coefficient $\hat{\beta}_1$ in a single-predictor regression model:

$$\text{s.e.}(\hat{\beta}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}$$

and

$$\text{s.e.}(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{\sum (x_i - \bar{x})^2}}.$$

The obtained t values of $\hat{\beta}_0$ and $\hat{\beta}_n$ (n>0) are then checked against the t distribution to retrieve their *p* values, which are then compared with the significance level to determine whether or not the predictors are significant.

The stepwise selection method starts with an equation containing no predictor variables, only a constant term. Then predictors enter the regression model roughly in the descending order of their correlations with the response variable. (The precise ordering depends on the correlations of the predictors with each other as well as with the response variable.) When each predictor enters the model, all the predictors in the model will be examined, and the variables that do not have significant contribution to the variability of the response variable will be dropped out of the model. The final model reached thus includes only the predictors that contribute significantly to the variability of the response variable.

Appendix B
CST 2004 Math Performance Level Cut Scores

Mathematics

Grade	Far Below Basic	Below Basic	Basic	Proficient	Advanced
2	150–235	236–299	300–349	350–413	414–600
3	150–235	236–299	300–349	350–413	414–600
4	150–244	245–299	300–349	350–400	401–600
5	150–247	248–299	300–349	350–429	430–600
6	150–252	253–299	300–349	350–414	415–600
7	150–256	257–299	300–349	350–413	414–600
General Mathematics*	150–256	257–299	300–349	350–413	414–600
Algebra I	150–252	253–299	300–349	350–427	428–600
Geometry	150–246	247–299	300–349	350–417	418–600

* All students in grades eight and nine who had not yet completed or were not enrolled in discipline specific, standards-based math courses or who were enrolled in the first year of a multi-year Algebra I course were required to take the General Mathematics CST. This test assesses the California Mathematics Standards for grades six and seven. Students in grades eight through eleven who had completed or were enrolled in discipline specific, standards-based math courses took California Mathematics Standards Tests in Algebra I, Geometry, Algebra II, or Integrated Mathematics 1, 2, or 3.

Appendix C
Correlation Tables of Analyzed Variables by Grade

Pearson Correlation Coefficients of Analyzed Variables -- Grade 6 (N = 68)

	Mean	*Std. Dev.*	CST 04 Math Scale Score	CST 03 Math Scale Score	Gender	Grad. of Entry	Is Black	Is Filip.	Is Hisp.	Is GATE	Eng. Flue.	Mean Q	Mean 1+ Ln(lat.)	Mean Nr. of Conc. / Strand / Stud.	Mean Nr. of Items / Conc. /Stud.	Mean Final Grade Plac./ Stud.
CST 2004 Math Scale Score	304.91	49.31	1.00	0.88	-0.26	-0.17	-0.50	0.44	-0.08	0.58	-0.23	-0.36	-0.62	-0.15	-0.69	0.29
CST 2003 Math Scale Score	300.02	65.93	0.88	1.00	-0.14	-0.20	-0.48	0.36	-0.13	0.53	-0.21	-0.41	-0.57	-0.19	-0.67	0.27
Gender	0.54	0.50	-0.26	-0.14	1.00	0.18	0.21	-0.26	0.01	-0.27	0.21	-0.05	0.42	0.11	0.25	-0.11
Grade of Entry	2.37	2.45	-0.17	-0.20	0.18	1.00	-0.13	0.06	0.24	-0.17	-0.18	-0.04	0.19	-0.06	0.11	-0.17
Is Black	0.45	0.50	-0.50	-0.48	0.21	-0.13	1.00	-0.35	-0.50	-0.31	0.48	0.15	0.49	0.14	0.31	-0.12
Is Filipino	0.13	0.34	0.44	0.36	-0.26	0.06	-0.35	1.00	-0.22	0.15	-0.23	-0.11	-0.24	0.13	-0.16	0.26
Is Hispanic	0.24	0.43	-0.08	-0.13	0.01	0.24	-0.50	-0.22	1.00	-0.08	-0.40	-0.10	-0.09	-0.13	-0.06	-0.16
Is GATE	0.10	0.31	0.58	0.53	-0.27	-0.17	-0.31	0.15	-0.08	1.00	0.02	-0.29	-0.50	-0.26	-0.50	0.23
English Fluency	2.53	0.91	-0.23	-0.21	0.21	-0.18	0.48	-0.23	-0.40	0.02	1.00	0.11	0.42	0.10	0.26	-0.06
Mean Q / Student	0.23	0.01	-0.36	-0.41	-0.05	-0.04	0.15	-0.11	-0.10	-0.29	0.11	1.00	0.14	-0.29	0.68	-0.35
Mean Ln(1+latency) / Student	2.88	0.28	-0.62	-0.57	0.42	0.19	0.49	-0.24	-0.09	-0.50	0.42	0.14	1.00	-0.02	0.62	-0.34
Mean Nr. of Concepts / Strand / Stud.	119.94	90.92	-0.15	-0.19	0.11	-0.06	0.14	0.13	-0.13	-0.26	0.10	-0.29	-0.02	1.00	-0.07	0.66
Mean Nr. of Items / Concepts / Stud.	6.27	0.99	-0.69	-0.67	0.25	0.11	0.31	-0.16	-0.06	-0.50	0.26	0.68	0.62	-0.07	1.00	-0.30
Mean Final Grade Placement / Stud.	0.47	0.05	0.29	0.27	-0.11	-0.17	-0.12	0.26	-0.16	0.23	-0.06	-0.35	-0.34	0.66	-0.30	1.00

* The demographic variable Is Other Asian is excluded because of its low frequency.

Appendix C
Variable Means, Standard Deviations and Correlation Coefficients

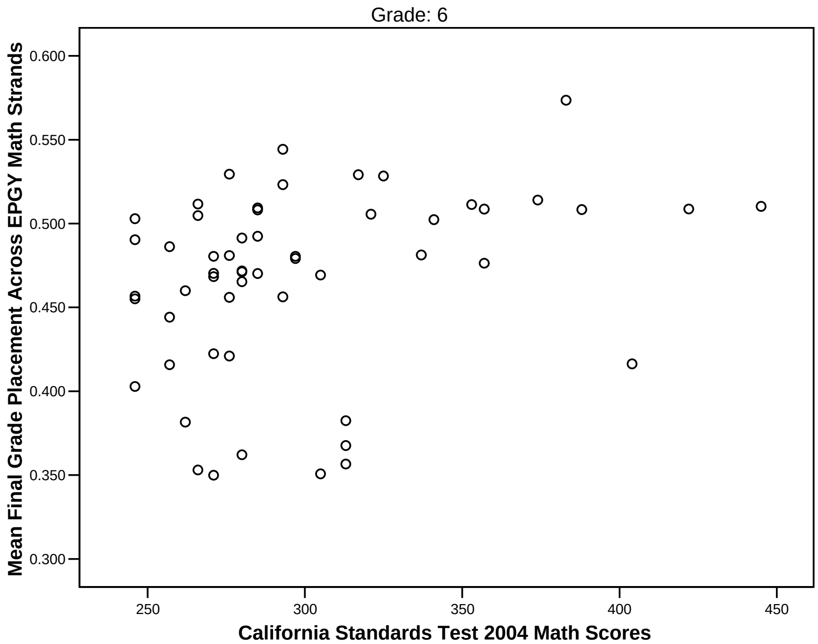
Pearson Correlation Coefficients of Analyzed Variables -- Grade 7 (N = 92)

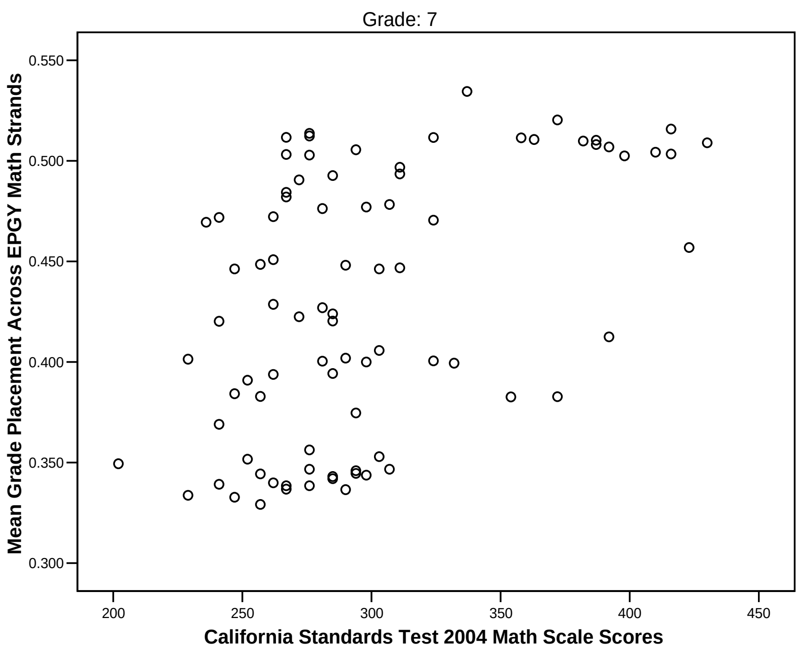
	Mean	*Std. Dev.*	CST 04 Math Scale Score	CST 03 Math Scale Score	Gender	Grad. of Entry	Is Black	Is Filip.	Is Hisp.	Is GATE	Eng. Flue.	Mean Q	Mean 1+ Ln(lat.)	Mean Nr. of Conc. / Strand / Stud.	Mean Nr. of Items / Conc. /Stud.	Mean Final Grade Plac./ Stud.
CST 2004 Math Scale Score	299.92	52.38	1.00	0.91	0.04	-0.12	-0.45	0.32	0.13	0.43	-0.17	-0.34	-0.57	0.31	-0.46	0.48
CST 2003 Math Scale Score	289.18	45.69	0.91	1.00	0.08	-0.14	-0.45	0.35	0.08	0.54	-0.21	-0.30	-0.47	0.26	-0.40	0.40
Gender	0.52	0.50	0.04	0.08	1.00	-0.16	0.00	0.12	-0.04	-0.06	-0.10	0.01	0.13	-0.08	0.10	-0.05
Grade of Entry	5.96	0.95	-0.12	-0.14	-0.16	1.00	-0.03	-0.01	0.05	0.01	-0.07	0.05	0.00	0.01	0.13	-0.13
Is Black	0.48	0.50	-0.45	-0.45	0.00	-0.03	1.00	-0.28	-0.63	-0.23	0.52	0.12	0.24	-0.19	0.10	-0.29
Is Filipino	0.08	0.27	0.32	0.35	0.12	-0.01	-0.28	1.00	-0.19	0.11	-0.04	-0.02	-0.39	0.19	-0.04	0.17
Is Hispanic	0.30	0.46	0.13	0.08	-0.04	0.05	-0.63	-0.19	1.00	0.05	-0.65	-0.10	0.04	0.00	-0.09	0.06
Is GATE	0.05	0.23	0.43	0.54	-0.06	0.01	-0.23	0.11	0.05	1.00	-0.16	-0.12	-0.20	-0.07	-0.20	0.23
English Fluency	2.43	0.98	-0.17	-0.21	-0.10	-0.07	0.52	-0.04	-0.65	-0.16	1.00	0.07	0.13	-0.07	0.10	-0.15
Mean Q / Student	0.23	0.02	-0.34	-0.30	0.01	0.05	0.12	-0.02	-0.10	-0.12	0.07	1.00	0.26	-0.28	0.80	-0.01
Mean Ln(1+latency) / Student	2.92	0.27	-0.57	-0.47	0.13	0.00	0.24	-0.39	0.04	-0.20	0.13	0.26	1.00	-0.40	0.38	-0.37
Mean Nr. of Concepts / Strand / Stud.	100.54	87.89	0.31	0.26	-0.08	0.01	-0.19	0.19	0.00	-0.07	-0.07	-0.28	-0.40	1.00	-0.29	0.60
Mean Nr. of Items / Concepts / Stud.	6.35	1.42	-0.46	-0.40	0.10	0.13	0.10	-0.04	-0.09	-0.20	0.10	0.80	0.38	-0.29	1.00	-0.16
Mean Final Grade Placement / Stud.	0.43	0.06	0.48	0.40	-0.05	-0.13	-0.29	0.17	0.06	0.23	-0.15	-0.01	-0.37	0.60	-0.16	1.00

* The demographic variable Is Other Asian is excluded because of its low frequency.

EPGY Curriculum Variables

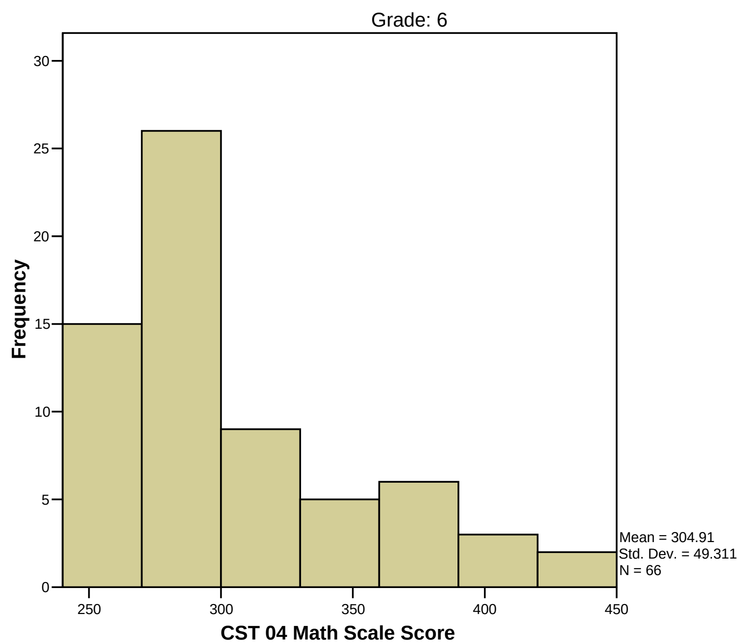
Appendix D
Scatter Plot of Mean Final Grade Placement Across EPGY Math Strands and California Standards Test 2004 Math Scores

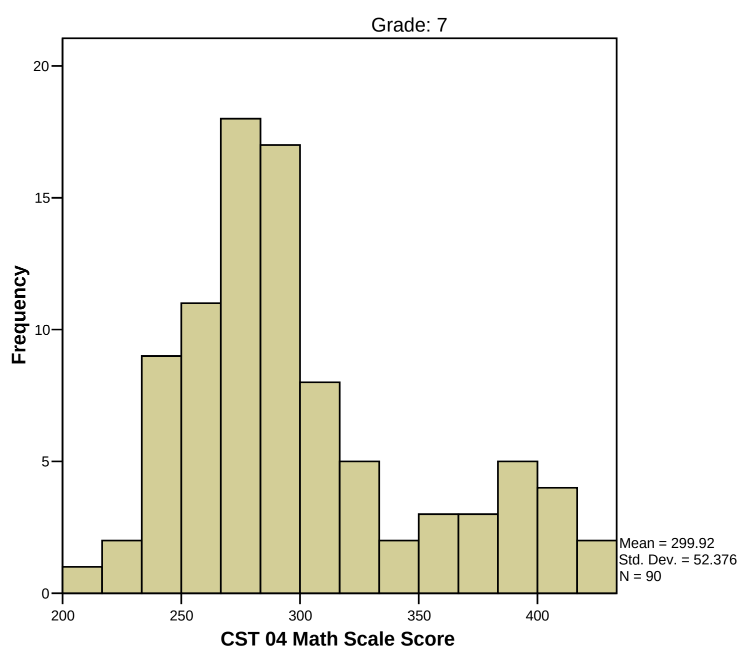




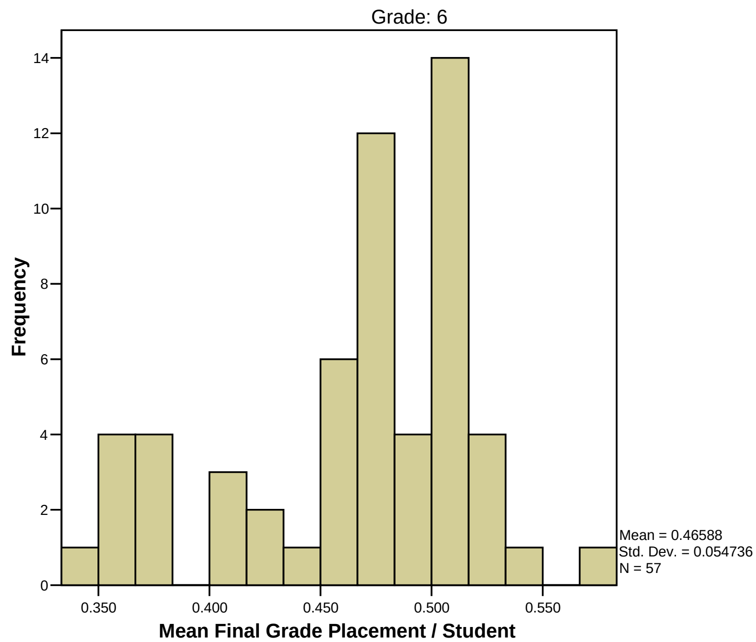
Appendix E
Histograms of Key Variables

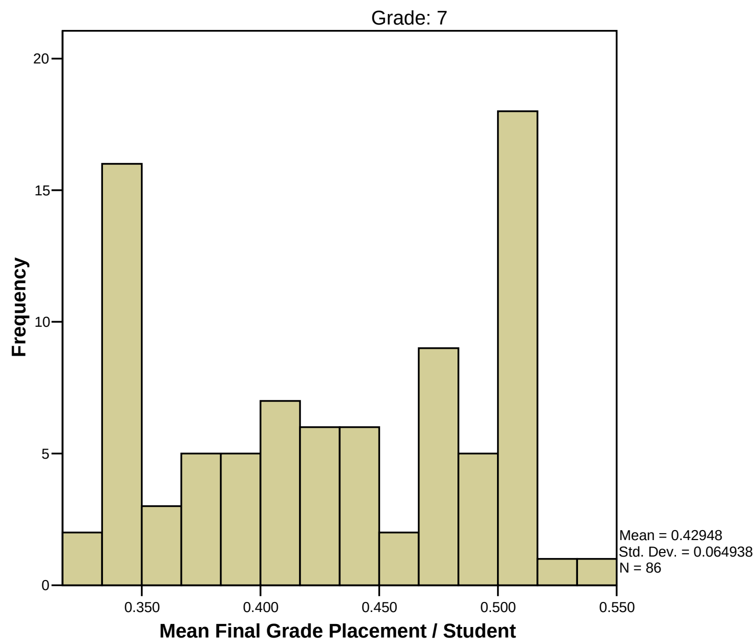
Histogram of Student Performances on California Standards Test 2004 Math





Histogram of Student Mean Final Grade Placement Across EPGY Strands





Histogram of Student Mean Ln(1+Latency)