Exploratory Quantitative Modeling of Curriculum Associates' i-Ready Student Instruction

Technical Report

Clint Sattler
Supervisor of Research and Evaluation
Knox County Schools
Department of Research, Evaluation, and Assessment

February 2020

Overview

Starting in the 2017-2018 school year (SY1718), the instructional leaders at Sarah Moore Greene Elementary (SMG) invested in a computer-based program to supplement teacher-led classroom instruction. SMG selected Curriculum Associates' i-Ready Student Instruction program to best integrate with their benchmark assessment suite. Students are assigned lesson modules in areas in RLA or math that have been identified as weaknesses on the benchmark assessments. Students are assessed at the end of each lesson to determine if they pass to lessons associated with other weaknesses. Students who fail to pass repeat the lesson. After multiple lesson failures, students are locked out of a lesson and teachers are alerted that the student is not making sufficient progress. The program vendor then suggests direct teacher intervention to better diagnose student needs.

The Knox County School's (KCS) Department of Research, Evaluation and Assessment (REA) analyzed data associated with i-Ready instruction at the end of SY1718. The study considered changes in normative measures on i-Ready benchmarks and the Tennessee state assessment as pertinent outcome variables. Curriculum Associates provided downloads of the data associated with i-Ready instruction usage to aid in the analysis. Because the study was constrained to one school, there was a limited number of student-level data points that could be used in the analysis. Due to concerns related to degrees of freedom, principal component analysis (PCA) was used to reduce the number of input variables. Principal component scores were used as input (predictor) variables in subsequent regression models. Regressions were limited to outcomes associated with math because there were very few students using the i-Ready RLA instruction modules.

The initial findings suggest that some linear combinations of i-Ready instruction usage variables correlated significantly with student growth on the i-Ready math benchmarks, but did not correlate with growth on the math state assessment. The following factors may have contributed to the null finding for the state assessment:

- The i-Ready benchmark utilizes a horizontal scale that spans grades K through 12 whereas the state assessment utilizes within-grade level scales. The i-Ready scale can theoretically be used to estimate off-grade level growth more precisely than the state scale.
- SMG students were regularly assigned lessons that were one or more grade levels lower than their enrolled grade, so any student growth could only be captured on a scale that extended below the students' enrolled grade.

REA staff visited with some members of the Curriculum Associates support and technical teams in April of 2018 to discuss the findings and possible paths forward in the analysis. The Curriculum Associates team had favorable views of the approach of the study but did suggest the addition of two input variables in future analyses (the time lag between student enrollment and their first access data of the i-Ready instruction modules and the number of i-Ready flags that could indicate spurious student-level data). The use of i-Ready instruction had expanded in the district, so the Curriculum Associates team also suggested that a more representative student population be used in future modeling efforts. The larger number of data points among the SY1819 cohort of i-Ready users would also allow the REA team to model variables as single inputs rather than solely rely on data reduction techniques.

This study seeks to better understand how i-Ready instructional usage correlates to changes in i-Ready benchmark normal curve equivalents (NCEs) and state assessment NCEs across a broad spectrum of users. The study is less concerned with discovering heuristics regarding i-Ready instruction usage (i.e. a student who spends X additional amount of time on i-Ready instruction should expect an impact of Y on the outcome variable of interest) and more concerned with determining if there is evidence to support the generally-accepted theory of action regarding use of the i-Ready instruction product. The general consensus among KCS users was that the following factors would contribute to larger positive gains on the i-Ready benchmarks and the state assessment.

- Longer total exposure to i-Ready instruction
- Longer time spent on i-Ready instruction per login event
- More frequent exposure to i-Ready instruction
- Earlier exposure to i-Ready instruction
- An increased number of lessons attempted
- An increased number of lesson quizzes passed
- A decreased number of quiz attempts before a lesson was passed.

The above bullet points constitute the KCS i-Ready instruction theory of action referenced throughout this document.

The findings of the current study indicate that some of the variables associated with i-Ready instruction usage correlate to gains on both the i-Ready benchmark and the state assessment in RLA and math. However, this study suggests that i-Ready instructional data may best serve as a leading indicator of gains in student content knowledge. There was no causal evidence to suggest that the outcomes were impacted by simply exposing a student to i-Ready instruction. It seems more likely that the i-Ready usage data projects from the same latent construct that projects onto benchmark and state assessments. This hypothesis suggests that many KCS users may want to reconsider how they monitor and react to the data associated with i-Ready instruction.

Methodology
The author of this research did not attempt to study any characteristics associated with the implementation of the i-Ready instruction product. No data was included in the quantitative analysis regarding the level of fidelity of the i-Ready instruction deployment nor was data collected related to the integration of i-Ready instruction data into instructional practice. The author feels that this approach is appropriate for this study since we do not wish to isolate the effect of the i-Ready instruction program on the output variables. We instead wish to study how inputs and outputs are related when constrained by our current level of program supports and variation of in-the-field practices.

Methodologically, REA is studying the levels of model congruency within and between subjects. Specifically, we are seeking to determine if variables that are strongly correlated ($\alpha=0.05$) with the outcome variables follow a unified theory of action. Congruency of parameter estimates requires that the sign of the parameter estimate associated with significant variables remains the same in the i-Ready and state test models. The REA modeling hypothesizes that the correlations between input and outcome variables will generally be stronger when the outcome variable is related to the i-Ready benchmark. We make this hypothesis since Curriculum Associates can ensure a strong (content and rigor) alignment between their products more easily than they can assure alignment with an outside entity, such as the Tennessee Department of Education. Therefore, it is unexpected that input variables would be significant predictors of state assessment results but would not be significant predictors of i-Ready benchmark results. The experience of REA staff (and a general consensus among research literature) is that math outcomes are longitudinally more stable. It would not be unexpected that the correlation between inputs and outputs in RLA would be less well-defined when compared to math results.

Exploratory hierarchical linear modeling (HLM) was initially deployed using an iterative approach to study the relationships between the assumed input variables and the outcome variables of interest. Null models were used to determine if data was hierarchical. Null models with an interclass correlation coefficient (ICC) of 10% or higher were deemed hierarchical in nature. We considered a three-level hierarchy in which students were fully nested under teachers and teachers were fully nested under schools. The HLM only contained random intercepts and did not include any group-level predictors. Input variables were introduced stepwise into the model and the Akaike Information Criterion (AIC) was monitored to determine if each variable correlated with significant variance in the outcome variables.

All HLM modeling was done in R (version 3.6.1) running on RStudio (version 1.2.1335) using the lme4 package (version 1.1.21). Parameter estimates were generated using maximum likelihood estimation. Restricted maximum likelihood estimation was not used due to the large number of groups in each model (112 teachers grouped under 16 schools for math and 69 teachers grouped under 12 schools for RLA). Confidence intervals were generated by the confint function in the stats package (version 3.6.1). 95% confidence intervals for fixed effects were determined using the bootstrap methodology with 500 replicates. Convergence tolerance was set to 0.0025.

The input variables considered in the analysis were:

- The school in which a student was enrolled at the time the state assessment was administered, entered as a level 3 factor variable
- The primary content-specific teacher linked to the student in the KCS schedule, entered as a level 2 factor variable
- X_G: An indicator of student grade-level at the end of SY1819, entered as an interval variable but considered a continuous variable
- X_1: A lagged measure of student achievement, entered as a continuous variable
    - For i-Ready benchmark, this was the Fall i-Ready benchmark NCE in a given subject
    - For the state assessment, this was the SY1718 state assessment NCE in a given subject
- X_2: An indicator of enrollment in special education (SPED) programming, entered as a dichotomous variable (a 1 signifying a student enrolled in a SPED program and a 0 indicating a student not enrolled in a SPED program)
- X_3: An indicator of socio-economic status, entered as a dichotomous variable (a 1 signifying a student classified as economically disadvantaged (ED) and a 0 indicating a student not classified as ED)
- X_4: An indicator of enrollment in English Language Learner (ELL) programming, entered as a dichotomous variable (a 1 signifying a student enrolled in an ELL program and a 0 indicating a student not enrolled in an ELL program)
- X_5: An indicator of student gender, entered as a factor variable (the reference factor was arbitrarily set as female)
- X_6: The primary race with which a student identified, entered as a factor variable (the reference factor was arbitrarily set as Native American)
- X_7: The overall number of i-Ready lessons passed, entered as an interval variable but treated as a continuous variable in modeling
- X_8: The overall i-Ready lesson pass rate, entered as a ratio variable but treated as a continuous variable in modeling. The pass rate was calculated as the number of lessons passed divided by the total number of lessons attempted.
- X_9: The overall time each student spent logged onto the i-Ready instruction platform (in hours), entered as a continuous variable

- X_{10}: The mean number of minutes a student spent on i-Ready instruction during each unique login event, entered as a continuous variable
- X_{11}: The mean number of days between student login events, entered as an interval variable but treated as a continuous variable
- X_{12}: The mean number of attempts a student made on a lesson quiz before passing, entered as a ratio variable but treated as a continuous variable
- X_{13}: The number of days that elapsed before the first day of school and the first time a student logged into the i-Ready instruction platform, entered as an interval variable but treated as a continuous variable

The number of i-Ready flags (which could indicate spurious student-level data) was not included in the final models. The inclusion of this variable led to convergence issues due to a low amount of between-student variance in this data. The author also originally considered including an input variable related to the number of unique lessons to which a student was exposed but high covariance with the number of lessons passed and the overall pass rate led to its eventual exclusion from the final models.

The i-Ready data was exported from the i-Ready website through the instructional usage reports. Student demographic data was exported from the Knox County student information system. Demographic data reflected the student demographics reported to the Tennessee Department of Education at the time of state testing. The modeling associated with changes in i-Ready benchmark and state test data only included data from students who had i-Ready instruction usage data. The results of this modeling cannot be used to assume any causal link between i-Ready instruction usage and these outcomes.

The results of the HLM were difficult to interpret due to within-and-between subject inconsistencies (see the Results section). It was hypothesized that the input variables were combining or interacting in meaningful ways that could not be captured when modeling inputs as single parameters. For example, a student spending large amounts of time using i-Ready instruction may be a student who is struggling with the content or the delivery platform or a student who is progressing through increasingly difficult lessons. Therefore, attempting to interpret the single parameter estimate associated with the total number of hours a student logged into i-Ready may not be enlightening. Additional screening models (not presented in this publication) were constructed using an interaction of each parameter with the lesson pass rate, but the additional complexity did not resolve the inconsistencies. The REA team decided to revisit the PCA methodology of the original SY1718 study to include linear combinations of correlated variables to reflect a more multi-dimensional character among the model predictors. PCA was accomplished using the R psych package (version 1.8.12). The varimax method was used to rotate the principal component axes. Variables that had a loading less than 0.4 were not considered members of the principal component.

Methodology: Change in i-Ready Benchmark

The i-Ready benchmark data was exported through the i-Ready website benchmark reports. National percentile ranks were converted to normal curve equivalents. The outcome variable of interest was the student-level change in NCE between the Fall and Spring benchmarks. Students were dropped from the analysis if they did not take a benchmark assessment within the first two months or the last two months of the academic year. Math modeling included 2,609 students. RLA modeling included 1,041 students.

The exploratory process ceded the following HLMs for math and RLA.

$$Y_{ijk} = \beta_{0jk} + \beta_1 X_{1i} + \cdots + \beta_n X_{ni} + r_{ijk}$$
$$\beta_{0jk} = \delta_{00k} + u_{0jk}$$
$$\delta_{00k} = \gamma_{000} + v_{00k}$$

Where Y_{ijk} is the change in i-Ready NCE (SY1819 Spring Benchmark NCE minus SY1819 Fall Benchmark NCE) of student i with teacher j in school k, β_{0jk} is the random intercept associated with teacher j in school k, β_1 through β_n are the fixed-effect parameter estimates associated with student-level predictors X_{1i} to X_{ni} respectively, r_{ijk} is the random error associated with student i with teacher j in school k, δ_{00k} is the random intercept associated with school k, μ_{ojk} is the random error term associated with teacher j in school k, γ_{000} is the grand mean change in NCE, and ν_{00k} is the random error term associated with school k.

Methodology: Change in State Assessment Data

TNReady data was available from the state test vendor (Questar). State percentile ranks were converted to normal curve equivalents using SAS conversion tables. The outcome variable of interest was the student-level change in NCE between the SY1718 and SY1819 state assessments. This outcome variable restricted which students could be included in the analysis. State test data was only available among third through fifth-grade students and SAS did not provide third-grade NCE conversions in a timeframe that would allow for inclusion in this study. Math modeling included 1,648 students. RLA modeling included 1,041 students.

The exploratory process ceded the following HLMs for math and RLA.

$$Y_{ijk} = \beta_{0jk} + \beta_1 X_{1i} + \cdots + \beta_n X_{ni} + r_{ijk}$$
$$\beta_{0jk} = \delta_{00k} + u_{0jk}$$
$$\delta_{00k} = \gamma_{000} + v_{00k}$$

Where Y_{ijk} is the change in state test NCE (SY1819 subject-specific state NCE minus SY1718 subject-specific state NCE) of student i with teacher j in school k, β_{0jk} is the random intercept associated with teacher j in school k, β_1 through β_n are the fixed-effect parameter estimates associated with student-level predictors X_{1i} to X_{ni} respectively, r_{ijk} is the random error associated with student i with teacher j in school k, δ_{00k} is the random intercept associated with school k, μ_{ojk} is the random error term associated with teacher j in school k, γ_{000} is the grand mean change in NCE and ν_{00k} is the random error term associated with school k.

Methodology: Interrupted Series Modeling

Data from a subset of three schools were examined to better estimate the causal impact of i-Ready instruction on state test results. Adrian Burnett Elementary, Christenberry Elementary, and Shannondale Elementary did not deploy i-Ready instruction to their students during SY1718 but did during SY1819. Monitoring longitudinal changes in state assessment between these two years may allow us to better estimate the impact i-Ready instruction has on changes in state test NCEs.

The modeling could only include data from the fifth-grade cohort during each academic year because of the need to calculate a change in NCE and the need to control for lagged achievement. Data from the fourth-grade cohorts could not be included since the Tennessee Department of Education did not release the state test scaled score-to-NCE conversion tables for third-grade results in both academic years. Math modeling included 243 students. RLA modeling included 212 students.

The data used in the analysis was further reduced based on student-teacher linkages. The KCS schedule data and i-Ready instruction data indicated that there were only three RLA and three math teachers who taught fifth-grade in both SY1718 and SY1819 and had students who used i-Ready instruction during SY1819. Only student data associated with these six teachers were used in subsequent regression models in order to limit variation in the parameter estimates due to longitudinal changes in instructors.

The student demographic variables that were identified as significant predictors in the state assessment model were retained in the interrupted series model. Individual i-Ready instructional usage variables were not included in the interrupted series models because of the inconsistencies in levels of significance and directions of parameter estimates found during the i-Ready benchmark and state test modeling. Because of the small number of teachers ($n<5$), HLM was not used for this modeling to avoid singular fits. Modeling was accomplished using a general linear model with teachers entered as factors. The academic year was included as a predictor variable to estimate the correlation between the year in which i-Ready instruction was deployed (SY1819) and the change in state NCE. The subject-specific predictor variables were those that were identified as significant in the previous models. The input variables considered in the analysis were:

- X_1: The SY1718 (lagged) state assessment NCE in a given subject, entered as a continuous variable
- X_2: An indicator of enrollment in special education (SPED) programming, entered as a dichotomous variable (a 1 signifying a student enrolled in a SPED program and a 0 indicating a student not enrolled in a SPED program)
- X_3: An indicator of socio-economic status, entered as a dichotomous variable (a 1 signifying a student classified as economically disadvantaged (ED) and a 0 indicating a student not classified as ED)
- X_4: An indicator of enrollment in English Language Learner (ELL) programming, entered as a dichotomous variable (a 1 signifying a student enrolled in an ELL program and a 0 indicating a student not enrolled in an ELL program)
- X_5: The primary race with which a student identified, entered as a factor variable (the reference factor was arbitrarily set as Native American)
- X_6: The subject-specific teacher, entered as a factor (the reference factor was arbitrarily set)
- X_7: The academic year, entered as a factor (the reference factor was set to SY1718)

The individual i-Ready usage variables (X_7- X_{13}) were not included because of concerns about how best to account for the mean time between logins and the number of days that elapsed between the first day of school and the first time the i-Ready lessons were accessed.

The general linear models used for each subject were:

$$y_{math,i} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \varepsilon_i$$
$$y_{RLA,i} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \varepsilon_i$$

Where y_i is the change in state test NCE (SY1819 subject-specific state NCE minus SY1718 subject-specific state NCE) of student i, β_0 is the grand mean change in NCE, β_1 through β_7 are the parameter estimates associated with student-level predictors X_{1i} to X_{7i} respectively and ε_i is the error associated with student i.

Results

The findings from both the math and RLA modeling lead to some uncertainty related to the original KCS theory of action for i-Ready instruction. The original theory of action hypothesized that increases in variables such as total time on task and mean time on task per lesson would be positively correlated with changes in both i-Ready benchmark and state test NCE. The theory of action additionally hypothesized that decreases in mean days between logins, mean lesson attempts prior to passing a lesson and the number of instructional days that elapsed between the start of the academic year and the initial exposure to i-Ready instruction would correlate with increases in i-Ready benchmark and state test NCE.

The results of the modeling could not largely support this theory of action. Some variables were significantly correlated with outcomes associated with one assessment, but not the other. Other variables were significantly correlated with the outcomes on both assessments, but the signs (positive or negative) of the parameter estimates were opposite for the two outcomes, or the sign was opposite of that expected by the KCS theory of action.

It is possible that the KCS theory of action was based on an incorrect hypothesis. It may be incorrect to assume that variables such as total time on task and mean time on task are primarily indicators of i-Ready instruction dosing, but rather may be indicators of broader student struggle against subject-specific content. There is some correlation between the inverse of the lesson pass rate and the total time on task that could support this interpretation.

Results: Change in i-Ready Benchmark, Math

The null model indicates that there is evidence of a hierarchical structure within the data (ICC=11.7%). The results suggest using a fully nested HLM (student-teacher-school) is appropriate.

All student demographic variables (other than grade-level) were significant predictors in the model during step-wise exploration. The final model suggests that gender may not be a significant predictor after the i-Ready instruction usage data is included in the model. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 1.

Table 1: i-Ready Benchmark Fixed Effects, Math (Single Variables)

			Confidence Interval	
Parameter (Fixed Effects)	Estimate	Std. Error	2.5%	97.5%
Intercept	15.85	5.73	4.28	27.44
X_1: Lagged achievement	-0.26	0.01	-0.28	-0.23
X_2: SPED	-3.28	0.63	-4.46	-2.14
X_3: ED	-1.04	0.44	-1.93	-0.14
X_4: ELL	-4.04	0.82	-5.53	-2.57
X_5: Gender = Male	0.03	0.37	-0.76	0.70
X_6: Asian	-3.08	4.81	-13.62	6.77
X_6: Black or African American	-6.20	4.64	-16.43	3.76
X_6: Hispanic/Latino	-3.43	4.66	-13.33	6.95
X_6: Native Hawaiian/Pac. Islander	-11.02	8.04	-27.27	5.83
X_6: White	-5.63	4.63	-15.58	4.32
X_7: # i-Ready lessons passed	0.16	0.03	0.11	0.21
X_8: i-Ready lesson pass rate	0.06	0.02	0.02	0.10
X_9: i-Ready time on task (hrs)	-0.04	0.06	-0.16	0.07
X_{10}: Mean time on task per login (mins)	-0.01	0.03	-0.07	0.05
X_{11}: Mean time between logins (days)	0.04	0.02	-0.01	0.09
X_{12}: Mean lesson attempts before pass	-4.18	1.67	-7.38	-0.71
X_{13}: Days between start of school and 1st lesson	-0.01	0.01	-0.03	0.01

The variance associated with the random effects is contained in Table 2.

Table 2: i-Ready Benchmark Random Effects, Math (Single Variable)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	12.04	3.47
School	Intercept	4.62	2.15
Residual		82.90	9.11

Analysis of the fixed-effect parameter estimates associated with the final model is largely consistent with the theory of action. The model suggests that as students pass more lessons and pass lessons more frequently the higher the change in their math i-Ready NCE. Additionally, students who require more attempts at a lesson before they pass tend to have lower changes in math NCEs. The other i-Ready instruction variables have little impact on change in i-Ready math NCE, though there is a near-significant positive correlation between the mean time between i-Ready instruction logins and change in benchmark NCE. However, the sign associated with this parameter suggests that students with longer mean time between login events do better on the i-Ready benchmark.

The residuals of the final model appear to be linear when plotted against predicted values, normally distributed, and have a mean near zero (-1.63e-15). The HLM including the i-Ready instruction variables accounts for 22.5% of the squared residual variance present in the null model. The model without the i-Ready instruction variables accounts for 14.6% of the squared residual variance present in the null model.

Bartlett's test indicates that the data has an appropriate amount of variance for PCA to be useful (Chi squared = 10040.56, df=21, p value=0). The Kaiser-Meyer-Olkin factor adequacy statistic (KMO) indicates that the data has a relatively low amount of common variance (KMO=0.50). The inflection point of the scree chart occurs between 3 and 4 principal components. The four principal components used in further modeling account for 88% of the total variance in the input data. The structure of the math principal components is available in Figure 1 and Table 3. The principal component scores were calculated for each student based on the PCA structure in Figure 1.

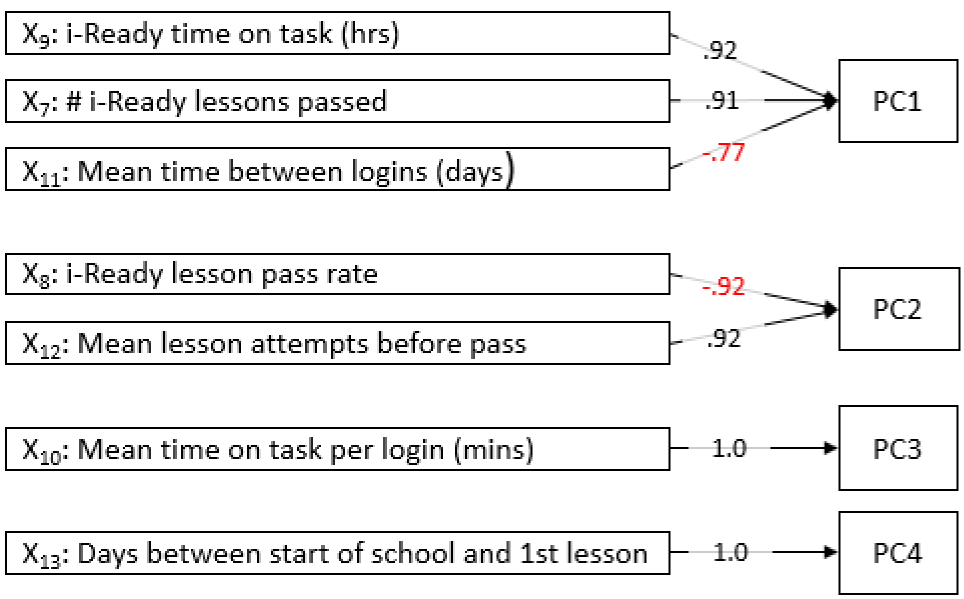

Figure 1: Math PCA Structure

Table 3: Math PCA Loadings and Variances

	PC1	PC2	PC3	PC4
SS loadings	2.33	1.71	1.09	1.04
Proportion Variance	0.33	0.24	0.16	0.15
Cumulative Variance	0.33	0.58	0.73	0.88
Proportion Explained	0.38	0.28	0.18	0.17
Cumulative Proportion	0.38	0.65	0.83	1

Based on the KCS theory of action, we would expect parameter estimates associated with PC1 to be positive and those associated with PC2 and PC4 to be negative. Ambiguity exists regarding the expected sign of the parameter estimates of PC3.

The principal component scores for each student were substituted for the unidimensional i-Ready instruction variables from the original Math i-Ready benchmark regression. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 4.

Table 4: i-Ready Benchmark Fixed Effects, Math (PCA Variables)

			Confidence Interval	
Parameter (Fixed Effects)	Estimate	Std. Error	2.5%	97.5%
Intercept	19.13	4.79	9.04	28.76
X_1: Lagged achievement	-0.25	0.01	-0.27	-0.22
X_2: SPED	-3.24	0.64	-4.60	-1.96
X_3: ED	-1.09	0.45	-1.88	-0.26
X_4: ELL	-3.97	0.83	-5.58	-2.29
X_5: Gender = Male	0.19	0.37	-0.53	0.90
X_6: Asian	-3.63	4.87	-13.33	5.90
X_6: Black or African American	-6.94	4.69	-15.94	2.67
X_6: Hispanic/Latino	-4.16	4.71	-13.36	5.42
X_6: Native Hawaiian/Pac. Islander	-12.20	8.13	-27.32	4.40
X_6: White	-6.25	4.68	-15.33	3.30
PC1	2.01	0.25	1.51	2.51
PC2	-2.08	0.20	-2.50	-1.68
PC3	-0.67	0.22	-1.11	-0.27
PC4	-0.90	0.22	-1.32	-0.43

The variance associated with the random effects is contained in Table 5.

Table 5: i-Ready Benchmark Random Effects, Math (PCA Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	12.02	3.47
School	Intercept	3.81	1.95
Residual		84.79	9.21

The HLM regression suggests that all principal components are significantly correlated with the change in i-Ready benchmark NCE. The sign (positive or negative) associated with principal components is in keeping with the KCS theory of action (with evidence suggesting that PC3 should be negatively correlated to change in NCE). The HLM including the i-Ready instruction variables accounts for 20.7% of the squared residual variance present in the null model. The model without the i-Ready instruction variables accounts for 14.6% of the squared residual variance present in the null model.

Results: Change in i-Ready Benchmark, RLA

The null model indicates that there is weak evidence of a hierarchical structure within the data (ICC=7.3%), but that both teacher and school effects are likely non-negligible. The RLA data was modeled using the same nested structure as the Math data for consistency. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 6.

Table 6: i-Ready Benchmark Fixed Effects, RLA (Single Variables)

Parameter (Fixed Effects)	Estimate	Std. Error	Confidence Interval 2.5%	Confidence Interval 97.5%
Intercept	4.29	7.19	-8.58	17.65
X_G: Student grade level	1.35	0.53	0.40	2.42
X_1: Lagged achievement	-0.20	0.02	-0.24	-0.17
X_2: SPED	-3.61	1.04	-5.62	-1.68
X_6: Asian	0.59	5.87	-10.13	11.17
X_6: Black or African American	-2.72	5.47	-12.66	7.88
X_6: Hispanic/Latino	-1.05	5.50	-10.60	9.26
X_6: White	-1.04	5.46	-10.49	9.27
X_7: # i-Ready lessons passed	0.05	0.05	-0.07	0.15
X_8: i-Ready lesson pass rate	0.04	0.02	0.001	0.09
X_9: i-Ready time on task (hrs)	0.08	0.09	-0.09	0.27
X_{10}: Mean time on task per login (mins)	-0.04	0.03	-0.11	0.03
X_{11}: Mean time between logins (days)	0.01	0.02	-0.03	0.05
X_{12}: Mean lesson attempts before pass	-1.40	2.06	-5.13	2.13
X_{13}: Days between start of school and 1st lesson	0.01	0.01	-0.01	0.04

Socio-economic status, ELL status, and gender were not significant predictors of changes in RLA i-Ready NCEs.

The variance associated with the random effects is contained in Table 7.

Table 7: i-Ready Benchmark Random Effects, RLA (Single Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	4.22	2.05
School	Intercept	6.15	2.48
Residual		86.11	9.28

There is only one i-Ready instruction parameter estimate for which the confidence interval does not include zero. This finding suggests that the correlation between the i-Ready instruction variables and the change in RLA i-Ready benchmark NCE is not as strong as the same correlation with the math data.

Analysis of the fixed-effect parameter estimates associated with the final model is largely consistent with the theory of action. The model provides some evidence that as students pass more lessons the higher the change in RLA i-Ready NCE. The number of school days that elapsed between the start of the academic year and the first i-Ready instruction lesson and the mean time on task have near significant correlations with the outcome variable.

The residuals of the final model appear to be linear when plotted against predicted values, normally distributed, and have a mean near zero (3.6e-15). The final regression model accounts for 15.3% of the squared residual variance present in the random (teacher-school) intercepts null model. The model without the i-Ready instruction variables accounts for 12.9% of the squared residual variance present in the null model.

Bartlett's test indicates that the data has an appropriate amount of variance for PCA to be useful (Chi squared = 3475.77, df=21, p value=0). The Kaiser-Meyer-Olkin factor adequacy statistic (KMO) indicates that the data has a relatively low amount of common variance (KMO=0.50). The inflection point of the scree chart occurs between 3 and 4 principal components. The four principal components used in further modeling account for 88% of the total variance in the input data. The structure of the RLA principal components is available in Figure 2 and Table 8. The principal component scores were calculated for each student based on the PCA structure in Figure 2.

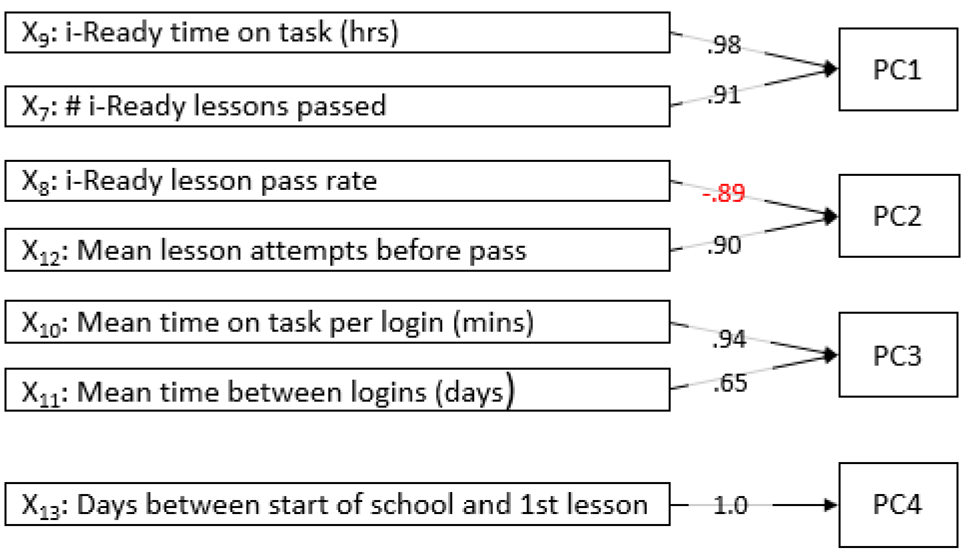

Figure 2: RLA PCA Structure

Table 8: RLA PCA Loadings and Variances

	PC1	PC2	PC3	PC4
SS loadings	2.04	1.61	1.5	1.01
Proportion Variance	0.29	0.23	0.21	0.14
Cumulative Variance	0.29	0.52	0.73	0.88
Proportion Explained	0.33	0.26	0.24	0.16
Cumulative Proportion	0.33	0.59	0.84	1

Based on the KCS theory of action, we would expect parameter estimates associated with PC1 to be positive and those associated with PC2 and PC4 to be negative. The elements that constitute each principal component are similar in math and RLA. The only observed difference was that X_{11} (mean days between logins) loaded on PC1 in the math PCA (with a negative weighting) but loads with a positive weighting on PC3 in the RLA PCA. The change in the sign of the weighting provides some evidence that the parameters estimates associated with PC3 should be negative (which aligns with the empirical finding in the math model, see Table 4).

The principal component scores for each student were substituted for the unidimensional i-Ready instruction variables from the original RLA i-Ready benchmark regression. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 9.

Table 9: i-Ready Benchmark Fixed Effects, RLA (PCA Variables)

			Confidence Interval	
Parameter (Fixed Effects)	Estimate	Std. Error	2.5%	97.5%
Intercept	6.78	6.01	-4.86	19.22
X_G: Student grade level	1.36	0.54	0.25	2.44
X_1: Lagged achievement	-0.20	0.02	-0.23	-0.17
X_2: SPED	-3.49	1.04	-5.43	-1.64
X_6: Asian	0.45	5.89	-11.56	12.98
X_6: Black or African American	-3.06	5.48	-14.85	7.57
X_6: Hispanic/Latino	-1.40	5.51	-13.09	9.20
X_6: White	-1.37	5.48	-12.67	9.66
PC1	0.81	0.35	0.06	1.52
PC2	-1.17	0.30	-1.77	-0.57
PC3	-0.67	0.31	-1.28	0.00
PC4	0.15	0.35	-0.59	0.87

The variance associated with the random effects is contained in Table 10.

Table 10: i-Ready Benchmark Random Effects, RLA (PCA Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	4.41	2.10
School	Intercept	5.22	2.28
Residual		86.61	9.31

The HLM regression suggests that PC1 and PC2 are significantly correlated with the change in i-Ready benchmark RLA NCE. The signs (positive or negative) associated with these principal components are in keeping with the KCS theory of action. PC3 was nearly significant and exhibited the expected negative correlation to the change in NCE. The residuals of the final model appear to be linear when plotted against predicted values, normally distributed, and have a mean near zero (6.71e-15). The final regression model accounts for 14.8% of the squared residual variance present in the random (teacher-school) intercepts null model. The model without the i-Ready instruction variables accounts for 12.9% of the squared residual variance present in the null model.

Results: Change in State Assessment Data, Math

The null model indicates that there is evidence of a hierarchical structure within the data (ICC=17.5%). The modeling results suggest that a 3-level HLM (student-teacher-school nesting) is appropriate. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 11.

Table 11: i-Ready-State Test Fixed Effects, Math (Single Variables)

			Confidence Interval	
Parameter (Fixed Effects)	Estimate	Std. Error	2.5%	97.5%
Intercept	-5.02	4.12	-12.97	2.66
X_1: Lagged achievement	-0.27	0.01	-0.29	-0.24
X_2: SPED	-4.93	0.81	-6.53	-3.34
X_3: ED	-1.43	0.57	-2.51	-0.35
X_4: ELL	-2.47	1.01	-4.60	-0.75
X_7: # i-Ready lessons passed	0.10	0.04	0.02	0.17
X_8: i-Ready lesson pass rate	0.09	0.03	0.04	0.14
X_9: i-Ready time on task (hrs)	0.04	0.07	-0.11	0.18
X_{10}: Mean time on task per login (mins)	0.09	0.04	0.02	0.16
X_{11}: Mean time between logins (days)	0.00	0.03	-0.06	0.05
X_{12}: Mean lesson attempts before pass	3.35	1.98	-0.70	7.29
X_{13}: Days between start of school and 1st lesson	0.02	0.01	0.00	0.04

The variance associated with the random effects is contained in Table 12.

Table 12: i-Ready-State Test Random Effects, Math (Single Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	21.48	4.64
School	Intercept	7.90	2.81
Residual		89.4	9.46

The model suggests that i-Ready Student Instruction usage data does correlate significantly with change in state Math NCE. The final regression model accounts for 17.9% of the squared residual variance present in the random (teacher-school) intercepts null model. The model without the i-Ready instruction variables accounts for 14.2% of the squared residual variance present in the null model.

The i-Ready benchmark parameter estimates associated with student demographics were similar in both direction and magnitude between the i-Ready benchmark and state math test models. The predictor variables associated with the number of i-Ready lessons a student passes and the student's lesson pass rate are significantly positively correlated with the outcome variable in both the i-Ready and state assessment models. However, this is not true for all of the other predictor variables. Specifically, the predictor variable associated with the mean number of minutes a student spends per lesson is significantly positively correlated with the change in state assessment NCE but not significantly correlated with the change in i-Ready Math benchmark NCE. Additionally, the input variable associated with the mean number of times a lesson is attempted is significantly negatively correlated with changes in the i-Ready benchmark but moderately positively correlated with change in state test NCE.

The principal component scores for each student were substituted for the unidimensional i-Ready instruction variables from the original math state assessment regression. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 13.

Table 13: i-Ready-State Test Fixed Effects, Math (PCA Variables)

			Confidence Interval	
Parameter (Fixed Effects)	Estimate	Std. Error	2.5%	97.5%
Intercept	12.07	1.32	9.48	14.61
X_1: Lagged achievement	-0.25	0.01	-0.27	-0.22
X_2: SPED	-4.75	0.81	-6.41	-3.17
X_3: ED	-1.42	0.58	-2.61	-0.27
X_4: ELL	-2.52	1.01	-4.47	-0.56
PC1	1.78	0.32	1.10	2.35
PC2	-0.87	0.26	-1.36	-0.33
PC3	0.44	0.26	-0.07	0.93
PC4	-0.13	0.31	-0.74	0.45

The variance associated with the random effects is contained in Table 14.

Table 14: i-Ready-State Test Random Effects, Math (PCA Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	21.00	4.58
School	Intercept	7.90	2.81
Residual		90.81	9.53

The HLM regression suggests that PC1 and PC2 are significantly correlated with the change in i-Ready benchmark NCE. The sign (positive or negative) associated with these two principal components is in accordance with the KCS theory of action with one exception. The sign associated with the PC3 parameter is opposite than expected by the theory of action, but this finding may be inconsequential since there is not a statistically significant correlation between PC3 and the change in math state test NCE. The residuals of the final model appear to be linear when plotted against predicted values, normally distributed, and have a mean near zero (-3.48e-15). The HLM including the i-Ready instruction variables accounts for 16.7% of the squared residual variance present in the null model. The model without the i-Ready instruction variables accounts for 14.2% of the squared residual variance present in the null model.

Results: Change in State Assessment Data, RLA

The null model provides weak evidence of a hierarchical structure within the data (ICC=4.0%). There is weak evidence to suggest that a 3 level HLM (student-teacher-school nesting) is appropriate in comparison to a non-hierarchical general linear model. For logical consistency, the analysis used a hierarchical structure even though the complexity of HLM may not be required by the data. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 15.

Table 15: i-Ready-State Test Fixed Effects, RLA (Single Variables)

Parameter (Fixed Effects)	Estimate	Std. Error	Confidence Interval 2.5%	Confidence Interval 97.5%
Intercept	37.25	13.10	11.00	62.01
X_1: Lagged achievement	-0.30	0.03	-0.35	-0.25
X_2: SPED	-5.54	1.64	-8.78	-2.31
X_6: Asian	-14.50	11.99	-39.16	10.20
X_6: Black or African American	-19.17	11.67	-41.69	4.70
X_6: Hispanic/Latino	-14.98	11.68	-37.75	9.48
X_6: White	-17.25	11.63	-39.33	6.74
X_7: # i-Ready lessons passed	0.21	0.09	0.04	0.38
X_8: i-Ready lesson pass rate	0.01	0.03	-0.05	0.06
X_9: i-Ready time on task (hrs)	-0.18	0.14	-0.45	0.08
X_{10}: Mean time on task per login (mins)	-0.01	0.05	-0.10	0.08
X_{11}: Mean time between logins (days)	0.02	0.03	-0.04	0.07
X_{12}: Mean lesson attempts before pass	-6.94	3.06	-12.52	-1.03
X_{13}: Days between start of school and 1st lesson	0.04	0.02	0.001	0.07

The variance associated with the random effects is contained in Table 16.

Table 16: i-Ready-State Test Random Effects, RLA (Single Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	4.27	2.07
School	Intercept	7.02	2.65
Residual		131.52	11.47

The model suggests that some variables associated with i-Ready instruction do correlate significantly with change in state RLA NCE. The residuals of the final model appear to be linear when plotted against predicted values, normally distributed, and have a mean near zero (-2.99e-15). The final regression model accounts for 20.0% of the squared residual variance present in the random intercepts (teacher-school) null model. The model without the i-Ready instruction variables accounts for 17.1% of the squared residual variance present in the null model.

The most significant predictors appear to be related to the overall number of i-Ready lessons a student passes, the mean number of times a student attempts a lesson, and the number of days that elapse before the start of the academic year and the first i-Ready instruction dosing.

The magnitude and the direction of the parameter estimate associated with the number of lessons passed is in accordance with the math model and the i-Ready theory of action. The association between the mean number of times a student attempts a lesson and the change in state test RLA NCE appears to be unique to the RLA dataset but is still logical under the i-Ready instruction theory of action. The association between the state test RLA NCE and the number of days that elapse before the start of the academic year and the first i-Ready instruction dosing does not adhere to the i-Ready theory of action. There is also one near-significant predictor in which the parameter estimate also does not follow the KCS theory of action. The model detects a weak negative correlation between the total length of time a student spends using the I-Ready instruction tool and their change in state test RLA NCE.

The i-Ready benchmark parameter estimates associated with student demographics were similar in both direction and magnitude than those of the state assessment model (with the exception of the parameter estimate associated with the Asian ethnic group). Variation between the models is most evident when considering which i-Ready instruction variables are significantly correlated with each outcome variable. The input variable associated with the students' lesson pass rate was the only significant predictor of change in RLA i-Ready NCE. This variable was not a significant predictor of the change in RLA NCE on the state assessment. The variables that were significantly correlated with the state test outcomes were not significant predictors in the i-Ready benchmark model.

The principal component scores for each student were substituted for the unidimensional i-Ready instruction variables from the original RLA state assessment regression. The fixed-effect parameter estimates and the bootstrapped confidence intervals for the parameter estimates can be found in Table 17.

Table 17: i-Ready-State Test Fixed Effects, RLA (PCA Variables)

Parameter (Fixed Effects)	Estimate	Std. Error	Confidence Interval 2.5%	Confidence Interval 97.5%
Intercept	29.66	11.83	6.46	54.34
X_1: Lagged achievement	-0.29	0.03	-0.35	-0.24
X_2: SPED	-5.16	1.64	-8.44	-1.90
X_6: Asian	-12.88	12.06	-37.01	11.25
X_6: Black or African American	-17.97	11.74	-41.47	5.80
X_6: Hispanic/Latino	-13.51	11.75	-36.54	10.40
X_6: White	-15.83	11.70	-39.59	8.15
PC1	0.04	0.55	-1.01	1.09
PC2	-2.05	0.47	-3.03	-1.04
PC3	-0.16	0.44	-0.98	0.68
PC4	0.83	0.49	-0.12	1.80

The variance associated with the random effects is contained in Table 18.

Table 18: i-Ready-State Test Random Effects, RLA (PCA Variables)

Groups	Name	Variance	Std.Dev.
Teacher:School	Intercept	5.18	2.28
School	Intercept	5.49	2.34
Residual		132.97	11.53

The HLM regression suggests that PC2 is significantly correlated with the change in state test NCE. The negative sign associated with this principal component is in keeping with the KCS theory of action. The signs associated with the PC1 and PC4 are opposite what was expected by the theory of action, but this may be inconsequential since there is not a statistically significant correlation between these variables and the change in RLA state test NCE. The final regression model accounts for 24.2% of the squared residual variance present in the random (teacher-school) intercepts null model. The final regression model accounts for 24.3% of the squared residual variance present in the random intercepts (teacher-school) null model. The model without the i-Ready instruction variables accounts for 22.2% of the squared residual variance present in the null model. The residuals are linear when plotted against predicted scores, normally distributed, and have a mean near zero (-9.71e-16).

Results: Interrupted Series Modeling, Math

As discussed in the methodology section, usage data for i-Ready (variables X_7-X_{13}) were not included in this phase of the analysis due to concerns about how to numerically model some of the data. The i-Ready usage statistics are available in Table 19 to orient the reader to possible difference between the students used in the benchmark and state test modeling and the interrupted time series math modeling. It may be important to note that the students included in the interrupted time series study generally accessed i-Ready later and spent less time with the product than the students included in the benchmark and state test models.

Table 19: Math i-Ready Usage Statistics

	Time Series			Benchmark/State Test Models		
Metric	Average	Std Dev	Median	Average	Std Dev	Median
Number of i-Ready lessons passed	12.70	7.84	12.00	29.51	20.61	25.00
i-Ready lesson pass rate	65.27	21.04	67.00	75.96	15.90	78.00
Total minutes on i-Ready instruction	588.09	269.86	532.10	1032.32	592.40	958.25
Mean minutes per login	29.68	7.54	28.97	27.13	9.19	25.40
Mean days between logins	12.57	9.11	9.65	9.58	9.93	6.79
Mean Lesson Attempts	1.30	0.18	1.29	1.23	0.17	1.20
Days between start of school and 1st lesson	74.54	45.25	85.00	44.92	31.31	36.00

The parameter estimates associated with the Math interrupted series model can be found in Table 20. The mean change in state math NCE in SY1819 was -0.47.

Table 20: i-Ready Time Series, Math

					Conf. Int.	
Parameter Estimates	Estimate	Std. Error	t stat.	p value	2.5%	97.5%
Intercept	1.10	3.08	0.36	0.72	-4.96	7.16
X_1: Lagged achievement	-0.04	0.05	-0.87	0.38	-0.14	0.05
X_2: SPED	1.77	2.76	0.64	0.52	-3.66	7.21
X_3: ED	-2.60	1.92	-1.36	0.18	-6.38	1.18
X_4: ELL	2.17	3.77	0.58	0.57	-5.26	9.59
X_6: Math Teacher B	5.54	1.95	2.84	0.00	1.70	9.39
X_6: Math Teacher C	8.77	2.29	3.83	0.00	4.26	13.28
X_7: Academic Year SY1819	-3.30	1.66	-1.98	0.05	-6.57	-0.02

The parameter estimate associated with the academic year suggests that the deployment of i-Ready instruction was significantly correlated with a negative change in state NCE. This estimate does not align with the KCS theory of action, but the direction of the parameter estimate is the same as the sign on the SY1819 change in state test NCE in this subject. The residuals appear to be normally distributed with a mean near zero (-6.4e-16) and linear when plotted against the predicted change in state NCE.

Results: Interrupted Series Modeling, RLA

As discussed in the methodology section, usage data for i-Ready (variables X_7-X_{13}) were not included in this phase of the analysis due to concerns about how to numerically model some of the data. The i-Ready usage statistics are available in Table 19 to orient the reader to possible difference between the students used in the benchmark and state test modeling and the interrupted time series RLA modeling. It may be important to note that the students included in the interrupted time series study generally accessed i-Ready later and spent less time with the product than the students included in the benchmark and state test models.

Table 21: RLA i-Ready Usage Statistics

	Time Series			Benchmark/State Test Models		
Metric	Average	Std Dev	Median	Average	Std Dev	Median
Number of i-Ready lessons passed	9.35	6.86	8.00	14.66	12.92	11.00
i-Ready lesson pass rate	62.20	24.04	64.50	23.43	18.63	19.00
Total minutes on i-Ready instruction	384.69	213.43	351.40	750.80	467.42	634.90
Mean minutes per login	23.56	10.04	21.71	31.92	14.42	28.43
Mean days between logins	15.95	13.40	12.17	16.42	19.93	11.00
Mean Lesson Attempts	1.27	0.23	1.23	1.26	0.20	1.24
Days between start of school and 1st lesson	88.91	50.69	93.00	48.28	34.84	35.00

The parameter estimates associated with the RLA interrupted series model can be found in Table 22. The mean change in state RLA NCE in SY1819 was 2.56.

Table 22: i-Ready Time Series, RLA

					Conf. Int.	
Parameter Estimates	Estimate	Std. Error	t stat.	p value	2.5%	97.5%
Intercept	20.16	7.60	2.65	0.01	5.18	35.15
X_1: Lagged achievement	-0.24	0.05	-4.90	0.00	-0.33	-0.14
X_2: SPED	-5.42	2.97	-1.83	0.07	-11.27	0.42
X_5: Black or African American	-8.93	7.62	-1.17	0.24	-23.96	6.10
X_5: Hispanic/Latino	-4.46	7.35	-0.61	0.54	-18.95	10.03
X_5: White	-7.54	7.17	-1.05	0.29	-21.68	6.60
X_6: RLA Teacher B	1.64	2.15	0.76	0.45	-2.60	5.88

X_6: RLA Teacher C	3.27	2.24	1.46	0.15	-1.14	7.68
X_7: Academic Year SY1819	1.30	1.76	0.74	0.46	-2.17	4.78

The parameter estimate associated with the academic year suggests that the deployment of i-Ready instruction was not significantly correlated with a change in state NCE. The significance of the estimate does not align with the KCS theory of action even though the sign of the estimate does align. The direction of the parameter estimate is the same as the sign on the SY1819 change in state test NCE in this subject. The residuals appear to be normally distributed with a mean near zero (7.2e-17) and linear when plotted against the predicted change in state NCE.

Conclusions & Considerations

The reader is reminded that this project was designed to study potential relationships between I-Ready instruction on student outcomes as constrained by the current quality of implementation in the Knox County Schools. The researcher is not interested in isolating an impact of the program itself, but rather estimating how the i-Ready instruction program is correlated to outcomes at the current level of implementation in KCS. This is not a trivial distinction.

This study provides some evidence to suggest that the variables associated with using i-Ready instruction are correlated with changes in benchmark and state test normal curve equivalents. i-Ready instruction usage data seems to describe small but significant sources of variance in the change in state assessment data (~2-4%). The magnitude of this effect seems reasonable given the time students spend engaged with the i-Ready instruction platform (in comparison to teacher-led instruction). However, the within-subject and between-subject results suggest that the relationship may not be a simple correlation between input and outcome variables. Regressions using linear combinations of i-Ready instruction usage provide results that are more consistent with the Knox County theory of action and REA modeling hypotheses and therefore preferred data models use principal component scores as inputs. The signs associated with statically significant ($\alpha=0.05$) i-Ready instruction variables are provided in Table 21. The PCA models exhibit fewer significant variables when the outcome is the state test and the signs associated with each parameter estimate are in accordance with the KCS i-Ready theory of action.

Table 23: Summary of Significant Findings

		Math		RLA	
Model Type	i-Ready Parameter	i-Ready Benchmark	State Test	i-Ready Benchmark	State Test
Single Parameter Model	X_7: # i-Ready lessons passed	+	+		+
	X_8: i-Ready lesson pass rate	+	+	+	
	X_9: i-Ready time on task (hrs)				
	X_{10}: Mean time on task per login (mins)		+		
	X_{11}: Mean time between logins (days)				
	X_{12}: Mean lesson attempts before pass	-			-
	X_{13}: Days start of school -> 1st lesson				+
PCA Model	PC1	+	+	+	
	PC2	-	-	-	-
	PC3	-			
	PC4	-			

These results suggest that heuristics regarding i-Ready instruction dosing and/or lesson quiz pass rates may not be adequate predictors of student gains on the state assessment. Better predictors appear to be linear combinations of i-Ready instruction variables; e.g. a linear combination of overall lesson quiz pass rate and the mean number of attempts a student makes at a lesson before passing (PC2 in both RLA and math models). These findings mirror recommendations from the Curriculum Associates support team, who suggest that students are most likely to be successful using i-Ready instruction who use the platform frequently and have a pass rate of at least 70%.

A different modeling methodology was deployed in an attempt to better determine if usage of i-Ready instruction had a casual impact on student outcomes as estimated on the state test. The results of this study caused the REA team to question the assumed directionality of the correlation between state test data and i-Ready instruction usage. The models that were generated as part of this study presupposed that i-Ready instruction data was an input variable that interacted with a latent variable to impact changes on the benchmark and state assessments (Figure 3).

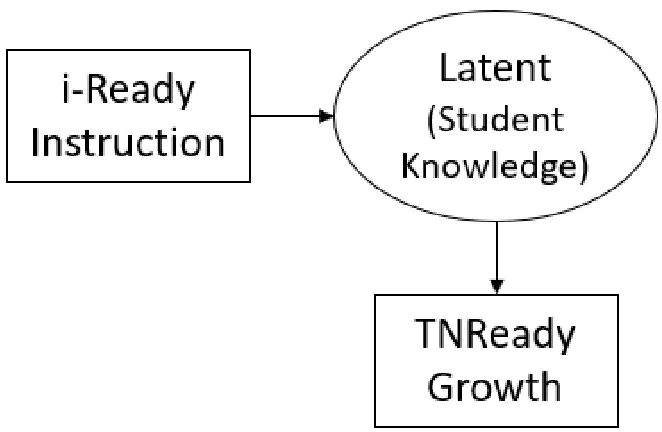

Figure 3: Pre-supposed i-Ready Instruction Theory of Action

Based on the findings of this study, REA now hypothesizes that i-Ready instruction data is more likely a projection of the same latent construct as that of TNReady (Figure 4).

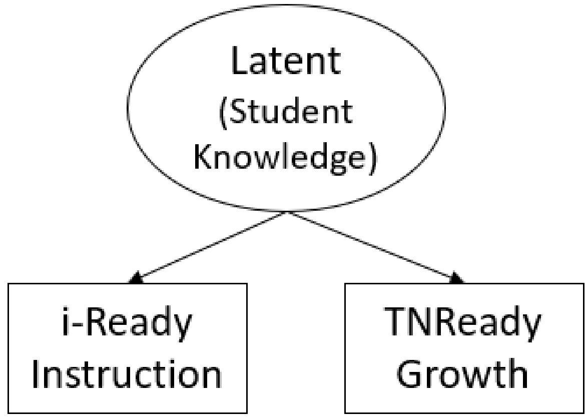

Figure 4: Current i-Ready Instruction Theory of Action

This new hypothesis may require some i-Ready instruction users to change the way they are deploying the product. It is proposed that i-Ready instruction data be viewed as a leading indicator of the changes in student content knowledge that will likely be reflected on the state assessment. Teachers can identify students who are struggling in their independent i-Ready instruction work, tailor a custom intervention for these students, then monitor the success of the intervention through changes in the i-Ready instruction data. For some students, the use of the online i-Ready instruction modules may be a sufficient intervention, but most other students may likely require direct instruction from a teacher. There was little evidence found in this study to suggest that using i-Ready instruction without a deliberate plan to use its output to inform instruction will lead to systematic gains in student performance on the state assessment.

The REA department will seek school-based partners with whom we can extend this study to make these findings actionable. Of primary interest will be constructing easy-to-use multi-dimensional monitoring tools for teachers to further test the proposed construct in Figure 3. The intent is to capture a parsimonious amount of i-Ready usage data to monitor growth in student content knowledge without tracking data associated with isolated variables. Recruitment efforts on these field trials will begin shortly after the publishing this report.

There are limitations to this study that may prevent the generalizability and reproducibility of the findings. The outcome variables for all of the models used normed data, which makes it impossible to understand what is the true counterfactual of this study. We cannot qualitatively describe how classroom instruction and the deployment of supplemental instructional resources vary across the norming population of the national i-Ready pool or the norming population of the state of Tennessee. Additionally, REA prefers to use hierarchical models (when possible) due to the assumed non-independence of results stemming from student-teacher-school linkages. This approach results in some low within-teacher sample sizes which may bias the models.

The need to control for a student's initial normal curve equivalent and some demographic variables when modeling changes in state test data was an unexpected finding. The SAS EVAAS manual makes it clear that their methodology does not require controlling for such variables, and this study uses techniques similar to the EVAAS model. It is hypothesized that the inclusion of repeated measures of student ability in the modeling would negate the need to use these variables. However, the lack of availability of third-grade NCEs meant that we could not include these repeated measures in our study.

Finally, the modeling methods used in this study may negatively impact the ability of future studies to reproduce these findings. Although the principal component models conform to the REA theory of action, this would not be true if the alpha value was extended to define significance at $\alpha=0.10$. Reproducing these findings over multiple cohorts would limit REA concerns over possible confirmation bias. However, deploying principal component analysis may hinder the ability to repeat the findings. Principal component analysis is a data reduction technique that seeks to explain the maximum amount of variance in the input variables in as few linear combinations as possible. PCA does not necessarily assume that these linear combinations are the result of an underlying latent structure. It is therefore possible that the models utilizing PCA in this study are over-fit to this specific dataset. The PCA structure identified in this study was similar to, but still different from, the PCA structure that was identified in the SY1718 study at Sarah Moore Greene. In the future, it may be advisable to employ a structural equation modeling (SEM) or factor analysis (FA) approach to better model the hypothesized latent constructs.