An Analysis of Outcomes Associated with the Year-long Reading Course

Technical Report

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

February 2019

Overview

Low literacy rates can have dramatic impacts on communities. Literacy rates have profound effects on physical and mental health, crime rates, and welfare dependency (Cree, 2012). Recent estimates place the economic cost of illiteracy at more than 300 billion US dollars (World Literacy Foundation, 2015). Because literacy has far-reaching impacts on society, Knox County Schools has continually made early literacy a priority during strategic planning (Knox County Schools, 2009; Knox County Schools, 2014; Knox County Schools, 2019).

Current research provides compelling evidence that reading is most effectively taught through methods that connect symbols to sounds, sounds to words, and words to meaning while engaging multiple linguistic and cognitive processors (Ehri, 2005; Foorman, 2016). Beginning readers must learn the relationships between letters and sounds to facilitate word decoding and encoding, but progressing readers must also learn to rapidly make meanings of words without translating words back to sounds (Castles, 2018). Students without these skills will be unable to comprehend text and will never attain true reading fluency (Lyon, 2002).

Although the research identifying how students learn to read is almost two decades old, educational institutions have been slow to adopt these findings into pedagogical practices and educational policy (National Reading Panel, 2000; Castles, 2018). In order to truly increase reading fluency, teachers require a deep understanding of student experiences, linguistic and cognitive abilities, and features of the texts (Moats, 1996). The Knox County Year-long Reading Course (YRC) was designed to help teachers better understand how children progress towards the comprehension of grade-level text and how teaching practices should reflect the science of reading.

The deployment strategy of the YRC has varied since its initial launch. The Knox County YRC began when district literacy coaches attended state training during the 2013-2014 school year (SY1314). The SY1314 YRC was largely experimental, and at the request of YRC designers, was not included as part of this study. Enrollment in the SY1415 YRC cohort was prioritized for teachers at schools with low reading scores on state assessments and teachers identified by school administrators as needing additional district support. The SY1516 cohort added teachers based on their (or their administrator's) expressed interest in the YRC concept. The focus of the YRC shifted to the inclusion of all elementary KCS English/Language Arts (ELA) teachers starting in SY1617.

Methodology: Personnel Allocation

Payroll records that connected substitute teacher allocations to YRC attendance were used to determine teacher participation in YRC classes. Student schedule data were extracted from the KCS data warehouse (EMIS) to link students to their ELA teachers. YRC field support logs were provided directly by the YRC support staff in order to track professional development activities.

Methodology: Teacher Concerns

The Stages of Concern Questionnaire was sent electronically to participants in all YRC cohorts. The questionnaire was developed by the Southwest Educational Laboratory as a framework to classify concerns about an initiative of innovation (Hall, 1977). The questionnaire was delivered to YRC participants via the Survey Monkey platform. Data for the SY1415 through SY1617 cohorts were collected between August and September of 2017. Pre-treatment data for the SY1718 YRC cohort were collected in the same time period. Post-treatment data were collected in May 2018. All data were collected anonymously.

Raw questionnaire scores were converted to percentile ranks per the Stages of Concern Questionnaire documentation. The greatest percentile rank for a given respondent was identified as their “peak” concern. Data were aggregated by cohort in order to identify frequencies and patterns in peak concerns.

Methodology: Outcomes

Ideally, a longitudinal analysis would be deployed to analyze the student outcome data associated with the YRC participants. This methodology would require the use of a consistent dataset from the years leading up to “treatment” (attending the YRC) and in subsequent years. However, there was no consistent student-level ELA data collected over the academic years of interest (SY1415-SY1718). The district has changed elementary ELA intervention screening assessments (from STAR Renaissance to the AIMSWeb suite of assessments) and the state ELA assessment was adjusted to assess more rigorous ELA standards (in SY1516). Therefore, estimates of program impact were generated from different quasi-experimental methods.

The analysis of outcome data used a matched-pair design between teachers who attended the YRC and a synthetic control group who did not receive the treatment. Coarsened Exact Matching (CEM) was used to create the synthetic control group. The input variables considered for treatment-to-control matching were teacher pay step (as a proxy for teacher experience), the percent of students in their classroom who were classified as economically disadvantaged (ED), lagged classroom-level mean performance on state exams, and lagged teacher observation scores. The exact variables used to create control-to-treatment matches varied according to the correlation between the input and outcome variables. Please consult the methodology section for each outcome variable for the complete details regarding the input variables used for treatment-to-control matching.

The CEM methodology requires the binning of input variables in order to create treatment-to-control matches. Sensitivity analysis of the bin intervals was accomplished by the inspection of variable imbalance statistics generated after CEM matching. The bin cut-points of input variables may vary depending on the level of correlation between the input and outcome. Input variables exhibiting high correlation to the outcomes were binned at tighter intervals than input variables exhibiting low correlation.

Two statistical tests were used to estimate the correlation between treatment and outcome after the CEM procedure. A fixed-effects general linear model (GLM) was used to model the relationship between the outcome variable of interest and a dichotomous dummy variable used to denote YRC treatment (1=treatment, 0=control, See Equation 1).

$$Outcome_i = \beta_{0i} + \beta_{1i} * Treatment\ Condition_i$$

Equation 1: GLM Model for YRC Analysis

The β_1 terms allowed us to estimate the impact of YRC participation on the mean of the outcome variable (when compared to the control group). Visual inspection of residual distributions was used to determine if linear modeling of the data was appropriate. Results from the GLM were suppressed if the residuals exhibited evidence of bias in the β_1 estimate.

The Kolmogorov-Smirnov test allowed us to determine if the distributions of treatment and control groups were likely to have come from the same parent distribution. The nonparametric Kolmogorov-Smirnov test allowed for the estimation of YRC impacts even when the data were non-linear. The Kolmogorov-Smirnov test is a conservative hypothesis test, as simulation studies have shown that the test can be insensitive to small differences in the tails of distributions (Babu, 2004).

Due to the size of the samples, we chose an alpha=0.05 for all tests. Data were aggregated according to the number of years that have elapsed since a teacher attended their first YRC. All calculations were completed on R version 3.4.3 running on R Studio Version 1.0.143. Coarsened Exact Matching used the “cem” package (April 12, 2018). General linear modeling and Kolmogorov-Smirnov testing used the base statistical package in R.

It is important to note that the creation of a wholly uncontaminated control group was impossible under the deployment strategy of the YRC. Teachers who participated in the YRC would, by design, expose untreated teachers to YRC concepts during collaborative lesson planning and professional learning communities (PLCs). All findings should be interpreted with the understanding that control teachers may be implementing YRC strategies without attending the YRC. Accordingly, the findings from this study cannot establish causal links between the outcome variables and the treatment conditions.

Methodology: Teacher Observation Scores

Observation data were extracted from the state database for teacher evaluation scores (TNCompass). Observation scores were calculated as the arithmetic means of scores on anywhere from 23 to 64 indicators (depending on teacher license status and past performance). Observation scores included scores in the planning, environment, instruction, and professionalism domains.

Matched treatment-to-control pairs were created from teachers with similar lagged observation scores (binned at 0.025 point increments). Lagged observation scores for treated teachers were the scores collected prior to the year of YRC treatment. The lagged observation scores for the comparison teachers were extracted from the same academic year as that of their treated counterpart. For example, if a teacher was treated in SY1516, their lagged observation score would come from SY1415. The control teacher paired with this treated teacher would have their "lagged" observation score extracted from SY1415 as well.

Exploratory analysis indicated that the percentage of ED students in a teacher's class and a teacher's pay step exhibited moderate correlation with observation scores. However, the results of subsequent modeling were relatively insensitive to the inclusion of the ED and pay step variables. There was evidence that the variation in observation scores related to the ED and pay step variables are subsumed by the lagged observation score. Both the ED and pay step variables were removed from the model for parsimony and to maximize the statistical power of subsequent testing.

It is important to note that there was no way to be certain that observation data were collected during an ELA lesson for teachers instructing multiple subjects. Although this may be a limitation of this analysis, there is no reason to suspect that the observation data collected in the treatment group was more or less likely to come from an ELA lesson when compared to the control group.

Methodology: AIMSWeb R-CBM

AIMSWeb Reading-Curriculum Based Measure (R-CBM) assessments were chosen to monitor outcomes related to reading fluency. Student R-CBM percentile rankings were converted to normal curve equivalents (NCEs). The arithmetic mean of the (student-level) spring benchmark R-CBM NCEs was linked to each ELA teacher for modeling. Nearly complete R-CBM datasets were available for SY1617 and SY1718 only. Data from earlier years were not used in the analysis. Spring benchmark R-CBM data were available for teachers in grades 1 through 5.

Matched treatment-to-control pairs were created from teachers with a similar percentage of ED students in their classroom (binned at 1% increments). All other variables were only weakly correlated with R-CBM outcomes.

Methodology: TCAP Writing Scores

Tennessee Comprehensive Assessment Program (TCAP) writing scores were chosen by the YRC staff as another outcome measure likely to be impacted by the YRC theory of action. TCAP sub-category scores related to writing focus and organization, writing development, language, and conventions were summed to calculate the raw TCAP writing score. The arithmetic means of student-level results were linked to each ELA teacher for modeling. TCAP writing data were only used from years SY1617 and SY1718 due to the shift in state ELA standards. TCAP writing data were only available for teachers in grades 3 through 5.

Matched treatment-to-control pairs were created from teachers with a similar percentage of ED students in their classroom (binned at 1% increments). All other variables were only weakly correlated with writing outcomes.

Results: Personnel Allocation

The current district cost associated with the YRC is relatively low. Prior to SY1718, three staff members were paid through general-purpose funds and their only responsibilities were associated with the YRC. However, in SY1718, these positions were allocated to school support for 50% of their time through the district microteams. As a result, it can be argued that the general-purpose budget is paying for 1.5 FTEs (Full Time Equivalent employees), in addition to course materials (books, printing, etc.). Pay for substitute teachers is required since the YRC occurs during the school day but these teachers are paid from federal Title II funds. Total program cost for the last two years is estimated at 140,000 general purpose dollars per year. Without the YRC, the support staff would likely be allocated for full-time school support, resulting in a total savings of approximately $38,000 against the KCS general purpose budget.

The largest district investment in the YRC may be the impact on high-quality instructional time as a trade-off to professional development. YRC participants who attend all classes miss a total of 5 full instructional days in their classroom.

The size of the YRC cohort will likely decline in the future. Figure 1 shows the percentage of general education ELA teachers in grades K through 5 that have attended at least one session of the YRC. The data for SY1819 is projected based on the beginning of the year rosters for the YRC.

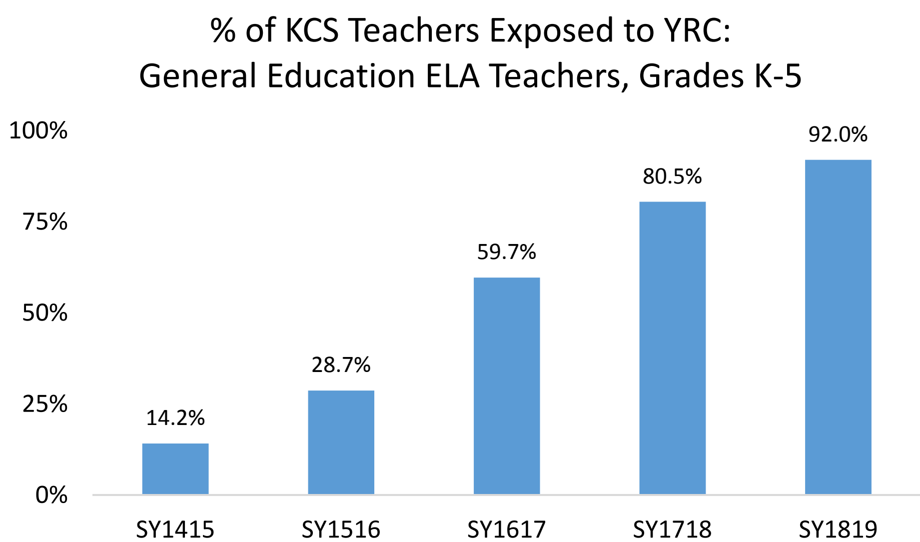

Figure 2: YRC Exposure Rates by Academic Year

At the end of SY1819, approximately 8% of K through 5 ELA teachers will not yet have attended at least one YRC class. However, employee turnover will necessitate more training if the goal is universal enrollment in the YRC. At the beginning of SY1819, 1,144 staff members were scheduled to teach at least one grade K through 5 ELA class. 245 of these teachers had not attended the YRC in previous years and 154 of these teachers were new to KCS. This suggests that KCS will have approximately 150 new ELA teachers in grades K through 5 each year due to employee attrition. This estimate does not include any ELA support staff (i.e. coaches), special education teachers, or English Language Learner (ELL) teachers who may enroll in the YRC. However, future YRC cohorts may include more of these staff members.

The YRC staff began systematically maintaining field support logs in SY1819. These logs provide a record of instructional support visits to schools. The provided support can be one-on-one with a teacher, support to a group of teachers, or direct support to a school's ELA instructional coach. The number of visits conducted by YRC staff through February of 2019 is available in Table 1.

Table 1: YRC Support Visit Frequency

Visit Type	N Visits	% of Total Visits
One-on-One Teacher Support	100	11.2%
Group Support	638	71.4%
Instructional Coach Support	156	17.4%

Results: Teacher Concerns

The response rates for the (post-treatment) Stages of Concern Questionnaire are contained in Table 2. The response rates were very low among the SY1415 and SY1516 cohorts. This may be due to the long interval between treatment and the administration of the questionnaire. Eleven (11) respondents did not provide their cohort year with their responses. All subsequent analyses assume that the respondents to the questionnaire constitute a representative sample of YRC alumni.

Table 2: Stages of Concern Response Rates

Cohort	N Sent Survey	N Responded to Survey	Response Rate
SY1415	220	18	8.2%
SY1516	526	33	6.3%
SY1617	250	91	36.4%
SY1718	251	47	18.7%

The distribution of peak concerns is contained in Figure 3. As evident from the plot, the majority of respondents had peaks corresponding to Stage 0 ("I am not concerned about it (the YRC)"). This profile provides evidence that some respondents would classify themselves as non-users of the YRC. The lack of a peak in Stage 1 ("I would like to know more about it") suggests that most non-use of the YRC innovation is not based on a lack of knowledge of the YRC content, but derives from other factors. Nearly forty-two percent of respondents (41.6%) to the survey indicated that they were in their first or second year of implementing some other major innovation or program, such as understanding new math standards or learning ELA standards to teach different grade levels.

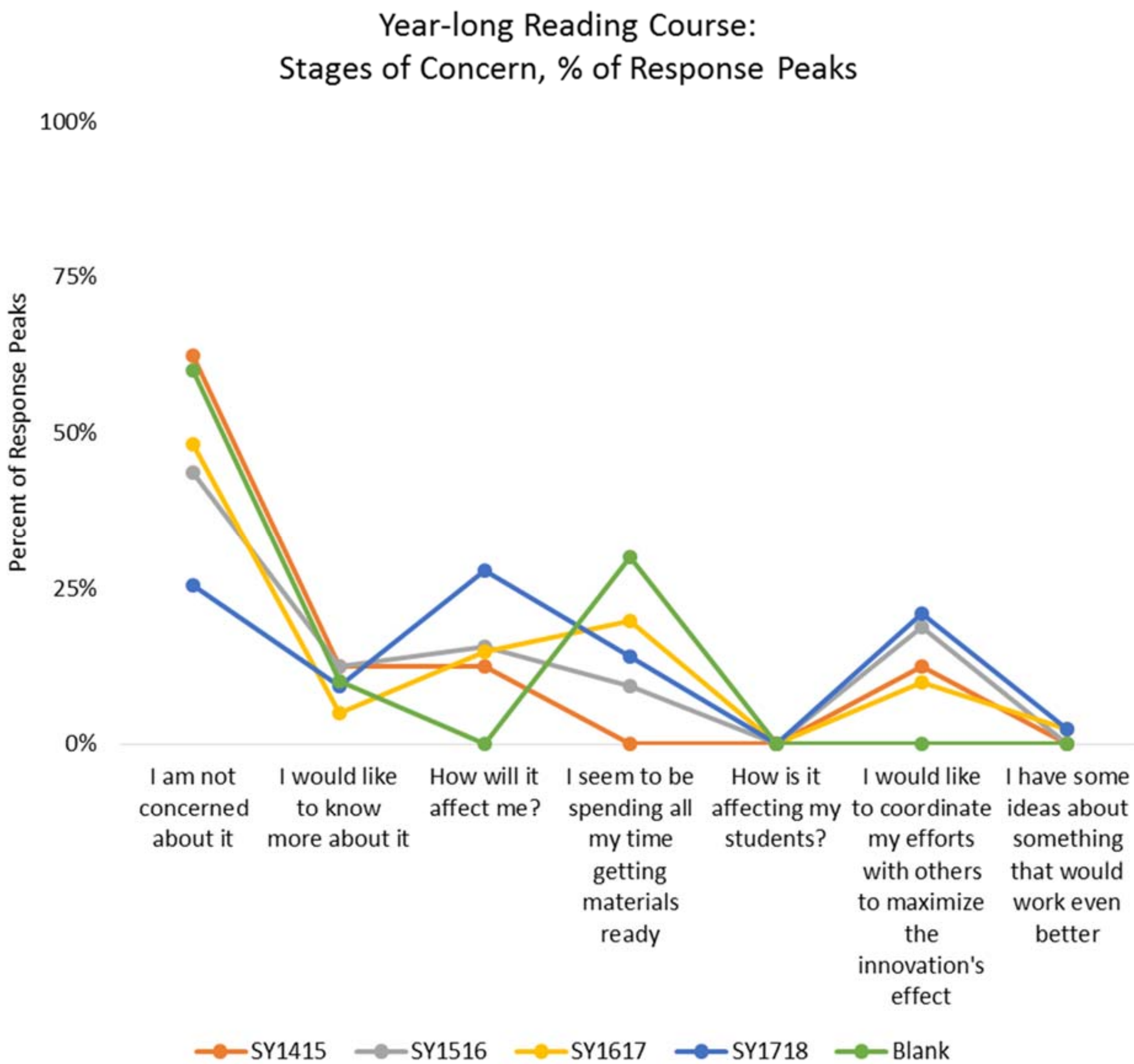

Figure 3: Response Peaks, Post-treatment Stages of Concern Questionnaire

There are also peaks visible at Stage 5 for each cohort ("I would like to coordinate my efforts with others to maximize the innovation's (YRC's) effect"). Peaks in this stage are common among team leaders who must coordinate work with others. It is possible that these peaks correspond to teachers looking to work within their professional learning communities (PLCs) to maximize the impact of the YRC. It is also possible that these peaks are artifacts of including instructional coaches and administrators in the data collection. Staff members in these roles would likely be concerned with the coordination of efforts within their building regarding the YRC. Unfortunately, respondents were not asked to provide their position when responding to the questionnaire.

The SY1718 cohort was sent the questionnaire in the Fall of SY1718 before YRC courses began. The same pool of respondents was sent questionnaires in the Spring of SY1718. The response rates for both administrations are contained in Table 3. The distribution of peak concerns is contained in Figure 4.

Table 3: SY1718 Cohort Pre/Post-Treatment Response Rates, Stages of Concern Questionnaire

Time Period	N Sent Survey	N Responded to Survey	Response Rate
Pre-Treatment	251	50	19.9%
Post-Treatment	251	47	18.7%

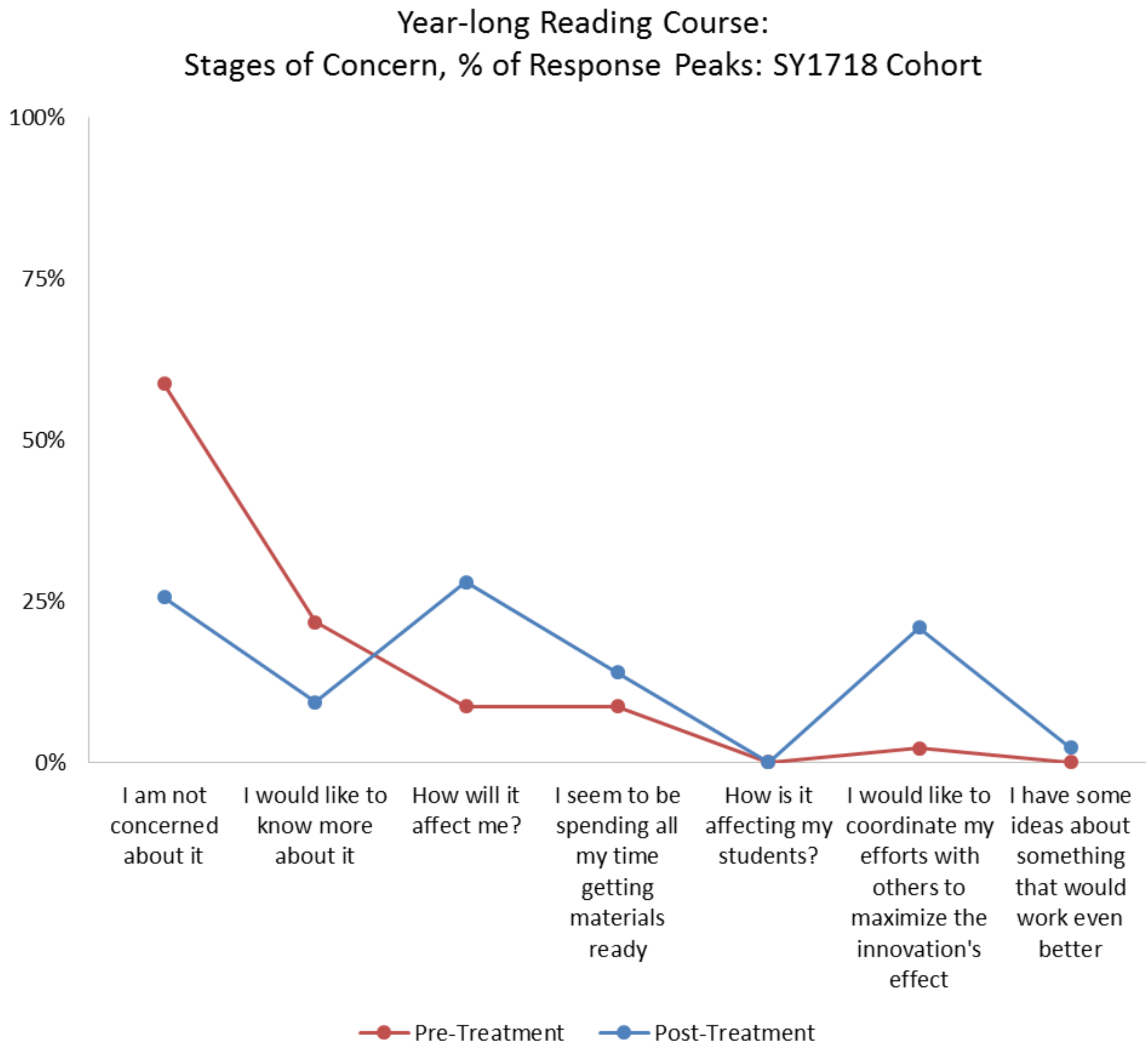

Figure 4: SY1718 Response Peaks, Pre/Post-Treatment, Stages of Concern Questionnaire

The downward shift in the Stage 0 concerns suggests that the respondents to the survey were much less likely to be classified as non-users after treatment. The shifting of the peaks to Stage 2 ("How will it affect me?") suggests that after attending the YRC, some teachers were concerned with how implementing the YRC would impact them personally and/or professionally. The literature suggests that these teachers would benefit from "non-threatening attempts… to discuss the innovation" (Hall, 1977). Respondents with these concerns are unlikely to consider the deep implementation of the YRC until their Stage 2 concerns are reduced.

The upward shift in Stage 5 shows that there is a pool of users ready to collaborate with others to maximize the impact of the YRC. Further evidence should be gathered to determine if these responses are associated with instructional support roles or with classroom instruction roles.

Results: Teacher Observation Scores

The results of the CEM matching are provided in Table 4. The results of the GLM fixed-effects modeling and the Kolmogorov-Smirnov testing are available in Table 5. Statistically significant results are denoted with an asterisk and bold text (α=0.05). The cumulative distribution functions for each teacher pool can be found in Figures 4 through 7.

Table 4: Teacher Observation Score CEM Matching Results

Teacher Pool	N in Each of Treatment and Control Groups	% of Treated Teachers with Control Group Match	Difference in Means (Obs. Score, Post CEM)
Year of Treatment	699	98.7%	3.00E-04
1 Year After Treatment	404	95.7%	4.95E-05
2 Years After Treatment	146	86.4%	8.90E-04
3 Years After Treatment	74	94.9%	5.41E-04

Table 5: Teacher Observation Score Hypothesis Testing Results. Significant results denoted with asterisk (α=0.05)

Teacher Pool	GLM Treatment Coefficient	GLM Treatment p value	Kolmogorov-Smirnov p value
Year of Treatment	-0.065	**0.011***	0.142
1 Year After Treatment	-0.082	**0.017***	**0.025***
2 Years After Treatment	-0.097	0.099	0.420
3 Years After Treatment	-0.032	0.700	0.644

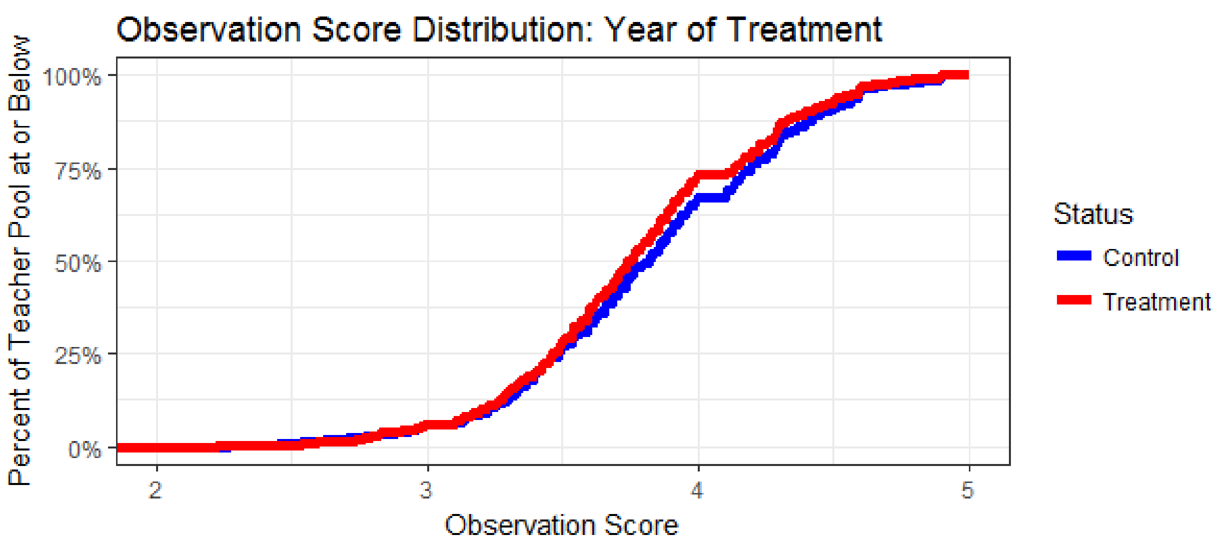

Figure 5: Teacher Observation Score CDF, Year of Treatment

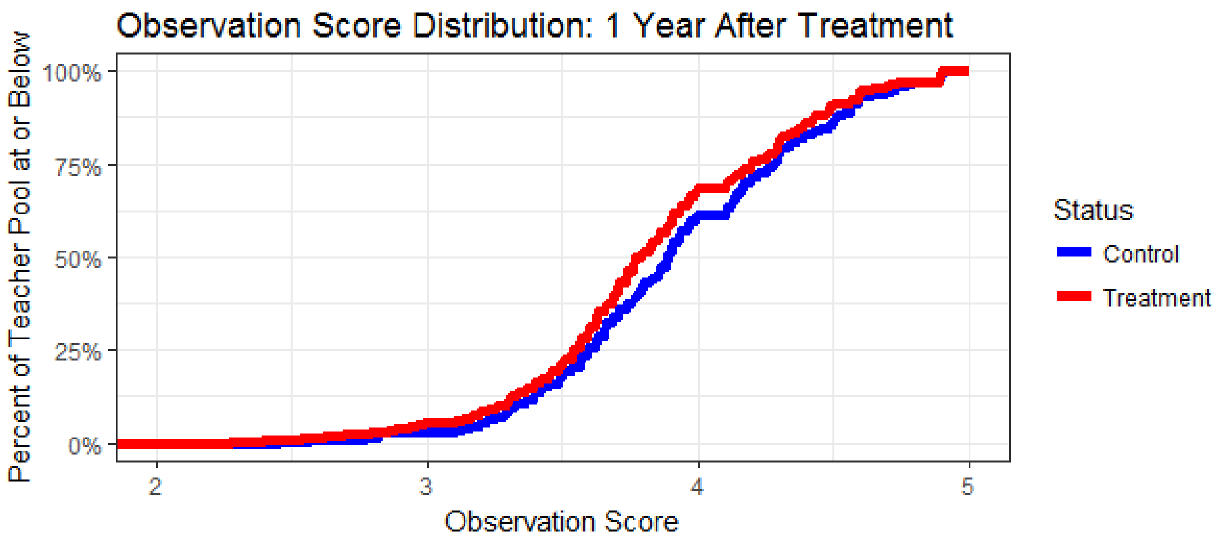

Figure 6: Teacher Observation Score CDF, 1 Year after Treatment

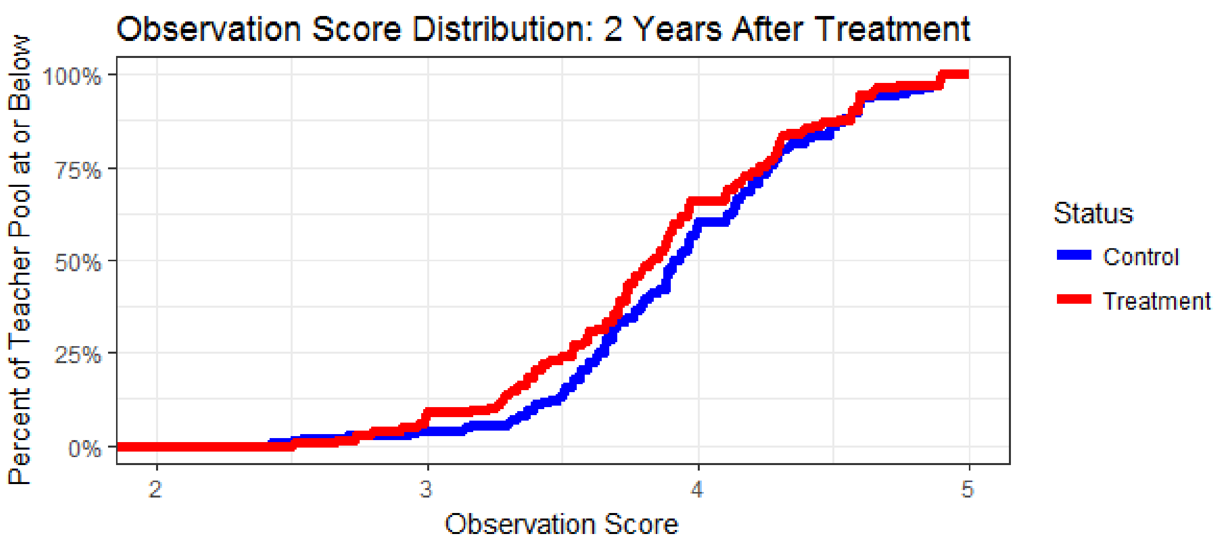

Figure 7: Teacher Observation Score CDF, 2 Years after Treatment

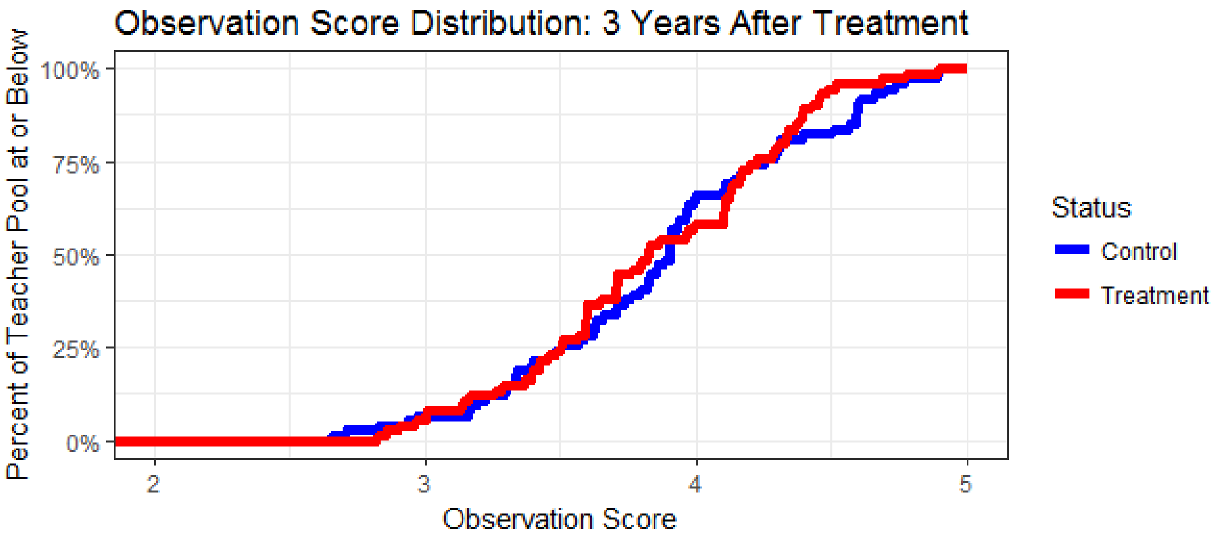

Figure 8: Teacher Observation Score CDF, 2 Years after Treatment

Using the results from the GLM model, we can reject the null hypothesis that the mean teacher observation score is no different between the following pools of teachers: teachers in the year in which they are enrolled in the YRC compared to their matched control pool, and teachers who are one year removed from the YRC when compared to their control pool. The sign of the coefficient on the treatment variable indicates that the mean observation score is lower for the YRC participants. We failed to reject the null hypothesis that the teacher observation scores were no different among teachers who were two and three years removed from their YRC enrollment compared to their respective control group. Directionally, the treatment coefficient for all teacher groups is negative.

Per the Kolmogorov-Smirnov test, we can reject the null hypothesis that the distribution of observation scores among control and treatment teachers come from the same distribution for teachers one year after treatment. We fail to reject the null hypothesis among the following teacher pools: Teachers in the year in which they are enrolled in the YRC, teachers who were two years removed from the YRC, and teachers that were three years removed from the YRC.

The CDFs in Figures 4 through 7 provide evidence that the differences in the distribution of observation scores tend to occur in the middle range of observation scores (from 3.5 to 4.5). The CDFs indicate that there is a similar distribution of teachers in both treatment and control groups scoring between 1.0 and 3.5 (although it is possible for teachers to have observation scores lower than 2.0, in practice no KCS elementary ELA teachers have observation scores below that threshold) and between 4.5 and 5.0.

The analyses provide some evidence that observation scores are lower among teachers in the year in which they are enrolled in the YRC and one year later when compared to teachers who did not attend the YRC. Although this finding may seem counterintuitive since the YRC should theoretically have positive impacts on instructional practice, it is possible that teachers struggle to put the complex YRC principles into practice. Observation scores may decrease in the short-term as teachers experiment with new instructional practices (Fullan, 2007). Anecdotally, the YRC instructors were concerned with how observers who had not attended the YRC would score YRC strategies on the observation rubric in the early years of implementation. This is less of a concern at present as most elementary observers have likely attended the YRC (See Figure 1).

Although these findings are statistically significant, they may not be practically significant. The largest estimated (directional) difference in observation score occurs among the pool of teachers that were two years removed from attending the YRC. In SY1718, the average indicator score was 3.88, with the average teacher being scored on 36 indicators. A deviation of -0.097 corresponds to a teacher scoring 1 point lower on 3 of the 36 indicators. This is equivalent to a teacher being scored one categorical level lower on 8.3% of the available indicators. It is uncertain that observers and teachers would even notice this level of variance in observation score.

Results: AIMSWeb R-CBM

The results of the CEM matching are provided in Table 6. The results of the GLM fixed-effects modeling and the Kolmogorov-Smirnov testing is available in Table 7. Statistically significant results are denoted with an asterisk and bold text (α=0.05). The cumulative distribution functions for each teacher pool can be found in Figures 8 through 11.

Table 6: Classroom-level Mean R-CBM NCE CEM Matching Results

Teacher Pool	N in Each of Treatment and Control Groups	% of Treated Teachers with Control Group Match	Difference in Means (ED, Post CEM)
Year of Treatment	373	71.7%	2.34E-04
1 Year After Treatment	328	73.7%	1.24E-04
2 Years After Treatment	198	78.0%	2.64E-04
3 Years After Treatment	99	92.5%	2.47E-04

Table 7: Classroom-level Mean R-CBM NCEs Hypothesis Testing Results. Significant results denoted with asterisk (α=0.05)

Teacher Pool	GLM Treatment Coefficient	GLM Treatment p value	Kolmogorov-Smirnov p value
Year of Treatment	4.2	**0.001***	**.022***
1 Year After Treatment	1.7	0.238	0.183
2 Years After Treatment	0.9	0.603	0.109

3 Years After Treatment	3.6	0.101	**.023***

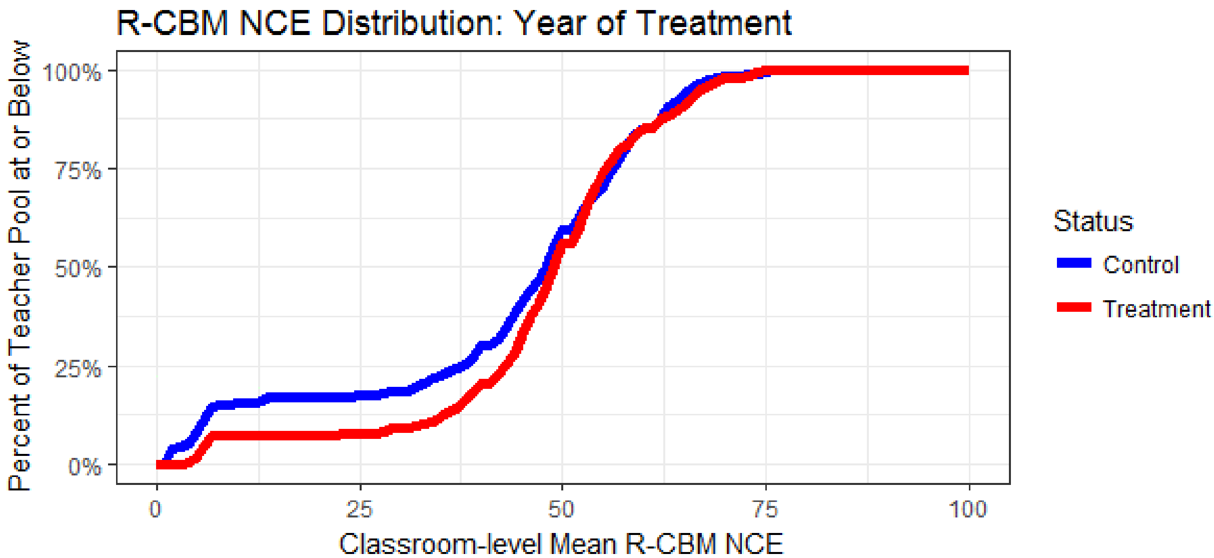

Figure 9: Classroom-level Mean R-CBM NCE, Year of Treatment

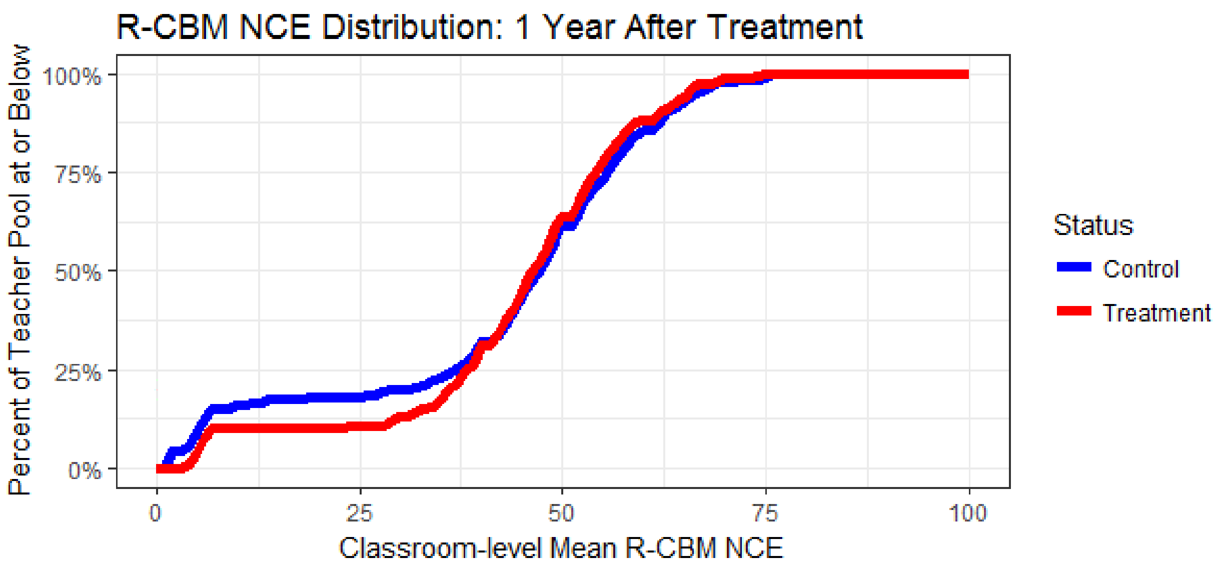

Figure 10: Classroom-level Mean R-CBM NCE, 1 Year after Treatment

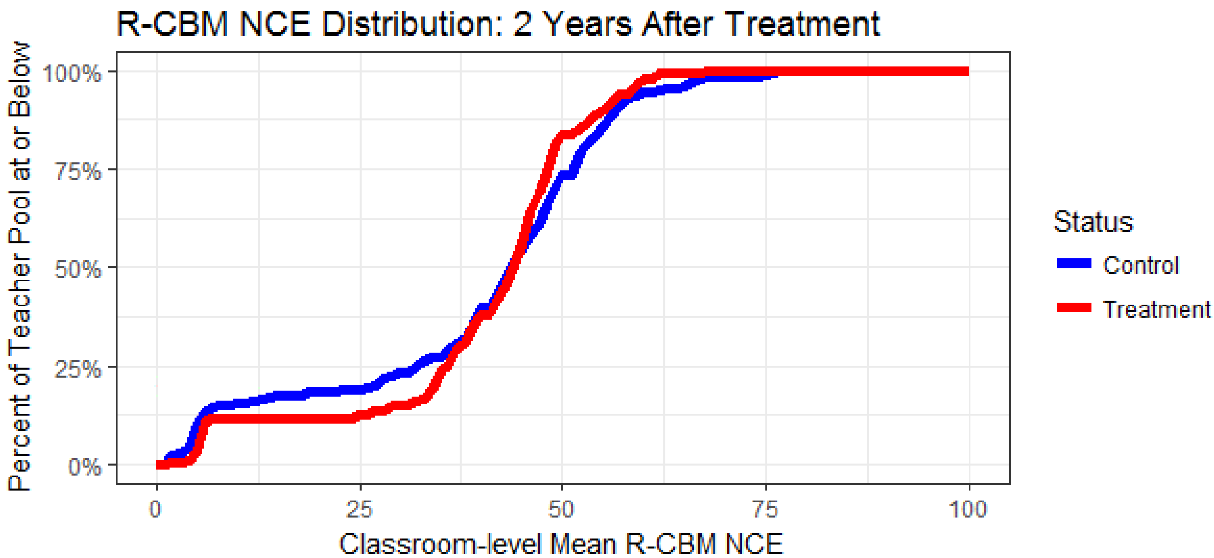

Figure 11: Classroom-level Mean R-CBM NCE, 2 Years after Treatment

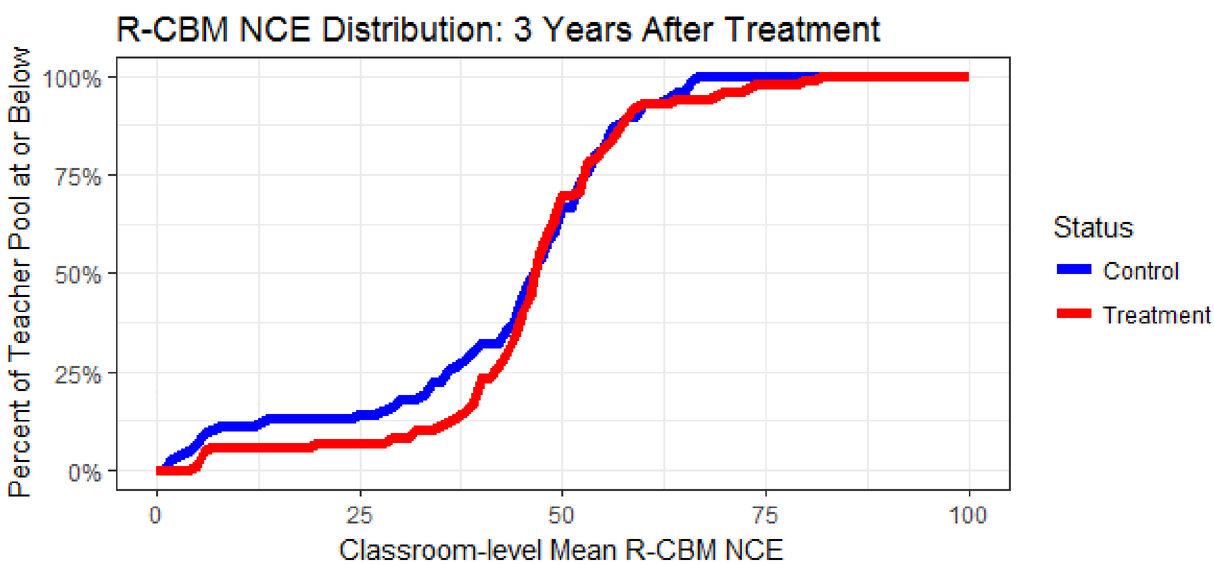

Figure 12: Classroom-level Mean R-CBM NCE, 3 Years after Treatment

We can reject the null hypothesis that the mean R-CBM NCE is no different between YRC teachers in the year in which they were enrolled in the YRC compared to their matched control pool. The sign of the coefficient on the treatment variable indicates that the mean R-CBM NCE is greater for the YRC participants. We failed to reject the null hypothesis that the mean R-CBM NCE was no different among teachers who were one, two, and three years removed from their YRC enrollment (when compared to the control teachers). Directionally, the treatment coefficient for all groups is positive.

We can reject the null hypothesis that the distribution of observation scores among control and treatment teachers come from the same distribution for teachers during the year of treatment and three years removed from treatment per the results of the Kolmogorov-Smirnov test. We fail to reject the null hypothesis among the following teacher pools: teachers who were one year removed from the YRC, and teachers that were two years removed from the YRC.

The CDFs in Figures 12 through 15 show that there were fewer YRC teachers with a mean R-CBM NCE less than 50 when compared to the control group. This is true for each of the four pools of teachers. It is interesting that the differences in the distributions do not materially extend beyond a mean R-CBM NCE of 50. This may be because teachers understand that moving students beyond this threshold may not materially impact reading comprehension (which is a tenet of the YRC). It is also possible that the reading fluency of lower performing students is more sensitive to YRC strategies, or that the NCE is impacted by ceiling effects of the measurement scale.

Methodology: TCAP Writing Scores

The results of the CEM matching are provided in Table 8. The results of the Kolmogorov-Smirnov testing is available in Table 9. The data were not sufficiently linear to generate unbiased results from the GLM model. Statistically significant results are denoted with an asterisk and bold text (α=0.05). The cumulative distribution functions for each teacher pool can be found in Figures 12 through 15.

Table 8: Classroom-level Mean Writing Scores CEM Matching Results

Teacher Pool	N in Each of Treatment and Control Groups	% of Treated Teachers with Control Group Match	Difference in Means (ED, Post CEM)
Year of Treatment	178	81.7%	8.61E-06
1 Year After Treatment	120	82.8%	3.28E-04
2 Years After Treatment	99	83.9%	1.50E-04
3 Years After Treatment	35	89.7%	5.41E-04

Table 9: Classroom-level Mean Writing Scores Hypothesis Testing Results. Significant results denoted with asterisk (α=0.05)

Teacher Pool	GLM Treatment Coefficient	GLM Treatment p value	Kolmogorov-Smirnov p value
Year of Treatment	-	-	**2.20E-16***
1 Year After Treatment	-	-	**8.26E-7***
2 Years After Treatment	-	-	**0.015***
3 Years After Treatment	-	-	**1.45E-4***

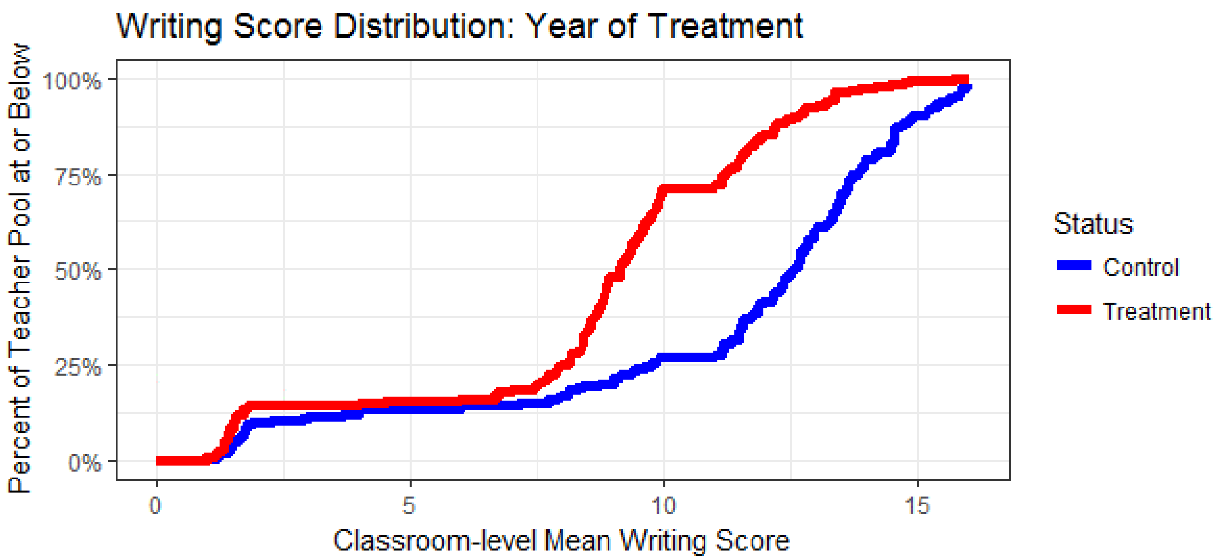

Figure 13: Classroom-level Mean Writing Scores CDF, Year of Treatment

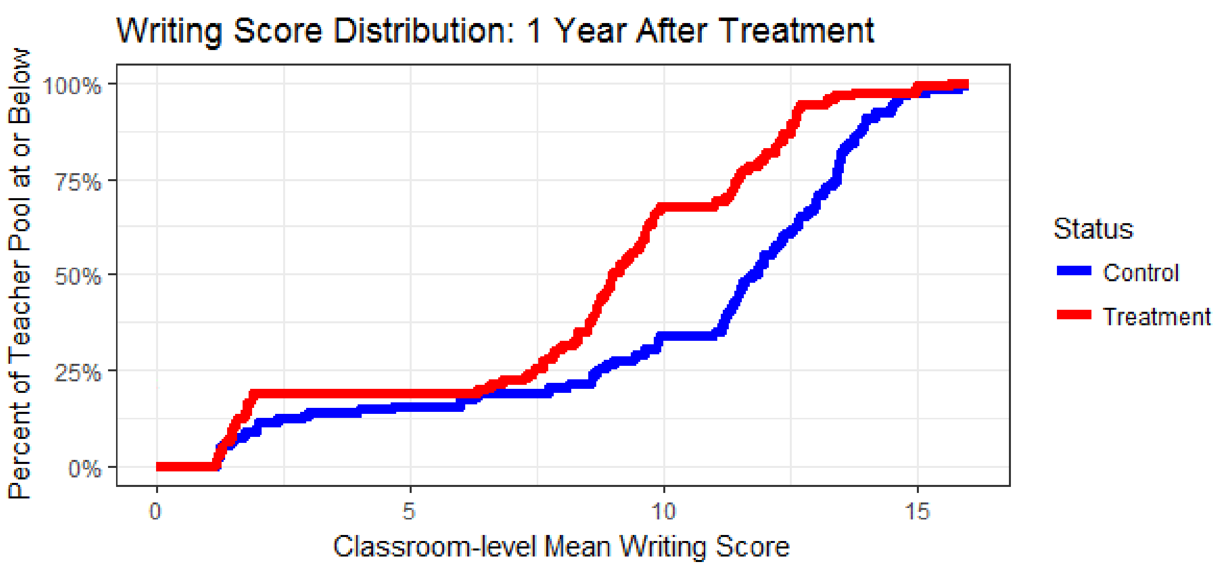

Figure 14: Classroom-level Mean Writing Scores CDF, 1 Year after Treatment

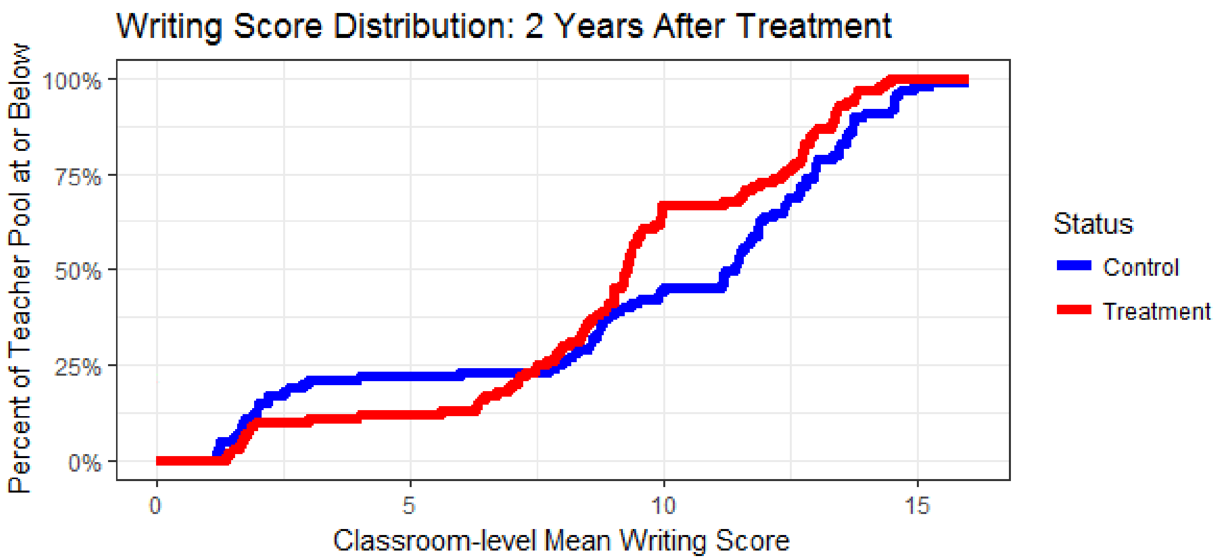

Figure 15: Classroom-level Mean Writing Scores CDF, 2 Years after Treatment

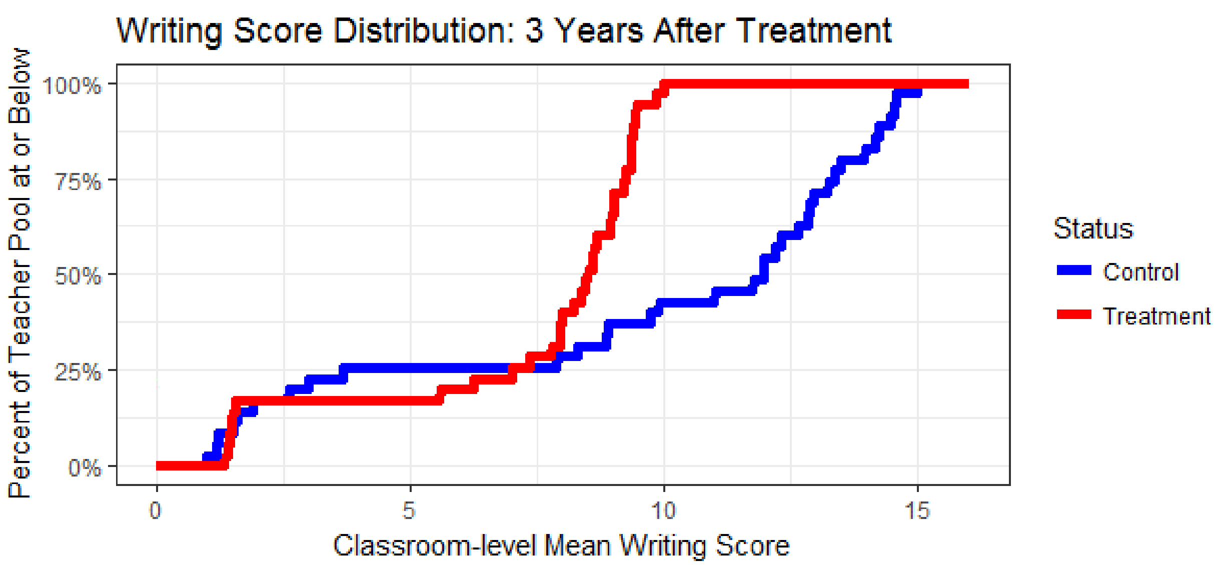

Figure 16: Classroom-level Mean Writing Scores CDF, 3 Years after Treatment

GLM modeling results are not presented for writing scores because the residuals were not normally distributed. The results of the Kolmogorov-Smirnov test indicate that we can reject the null hypothesis that the distribution of mean writing scores of teachers in the treatment and control groups came from the same population for all teacher pools.

Among teachers in the year in which they attended the YRC and teachers one year removed from the YRC, the percentage of teachers with a classroom-level mean writing score between approximately 1 and 7 is relatively equal. However, the CDFs suggest that fewer treatment teachers had lower mean writing scores (between 1 and 7) two and three years removed from treatment. The difference in the distributions that is being detected by the Kolmogorov-Smirnov test occurs beyond this threshold, indicating that there is a larger proportion of teachers with higher writing scores in the control group when compared to the treatment group.

This finding is somewhat surprising when viewed through the lens of the YRC logic model. The YRC promotes student writing as an essential component of reading lessons, and writing prompts created by the YRC support staff are typically used for assessment (both guided writing and "cold" writing). It is possible that writing is such a cognitively complex process that more time with the YRC strategies will be required before student writing scores are favorably impacted across all student ability levels.

Conclusions & Considerations

The current literature provides a strong rationale for the existence of a program like the Knox County Year-long Reading Course. The literature indicates that teacher preparation programs rarely expose pre-service teachers to the science of reading (Binks-Cantrell, 2012). The current YRC deployment strategy allows the district to expose our teaching staff to these concepts at a relatively low cost.

The data suggest that teachers generally fall into two distinct populations after participation in the YRC (assuming that the qualitative data collected as part of this study comes from a representative sample). There seems to be a subset of teachers who understand the YRC content, deploy the content in their classrooms, and are actively seeking partners with whom they can collaborate to add depth to their instruction. These teachers are likely confident in their knowledge of reading pedagogy, classroom management, and feel little burden from competing district initiatives. The key to a more impactful implementation of the YRC among these teachers may be to provide collaborative field support: a sounding board upon which these teachers can discuss ways to increase the depth of their instruction.

However, there is also evidence of a group of YRC alumni with very different concerns. The pre/post-treatment qualitative data suggest that a large population of teachers have concerns about how the implementation of the YRC will impact them personally and professionally after they matriculate from the YRC. These concerns may intensify as objective measures of success (such as observation scores) are impacted by "implementation dips" as instructional practice changes (Fullan, 2007). This may intensify personal and professional concerns about the YRC and make them less likely to implement YRC practices in their classrooms.

There is little evidence in the available data to suggest that teachers are concerned about a lack of knowledge of the YRC content, the impact of YRC implementation on students, or are concerned about competing reading instruction programs. This may be an indicator that teachers are generally aware of the YRC content and generally agree that the pedagogical approach presented in the YRC is the best path forward for their students. Previous KCS studies on the YRC provide additional qualitative evidence to strengthen this hypothesis (Abdelrazek, 2016).

Based on these data, YRC field support may consider taking a two-pronged approach to supporting YRC alumni in the field. One prong of support could focus on providing collaborative support to high-performing YRC implementers. However, the district should not ignore the concerns of teachers with lingering professional and personal concerns after attending the YRC. The second prong of field support could be tailored to alleviate these professional and personal concerns in an open and non-threatening way. Discussing advanced principles of the YRC among teachers who are still harboring personal and professional concerns may drive them towards the non-user end of the implementation spectrum. Addressing their concerns directly will be more likely to move these teachers to deeper levels of implementation.

It should be noted that the YRC staff was already aware of these personal and professional concerns. Early YRC courses were focused mainly around the hard science of how students learn to read. The YRC staff received feedback in which teachers expressed a desire for more explicit instruction regarding classroom implementation. As a result, the YRC has continually evolved to include increasing amounts of explicit instructional supports. YRC-created supports, such as lesson plans, writing prompts, and exit tickets are currently available to all district staff. YRC staff may want to monitor how these supports are used. Literature indicates that it is important the teachers have a deep understanding of how instructional strategies and instructional materials are connected to comprehending grade-level texts (Moats, 2009). Instructional staff at all levels of the district should work to ensure that these resources are not being deployed in a strictly procedural manner. Questions remain about how best this could be accomplished and monitored.

The student-level outcomes associated with the YRC exhibit promising (directional) returns when focused on the foundational literacy skill, reading fluency. The distributions of (class-level) mean AIMSWeb R-CBM NCEs seem to indicate that YRC alumni are more likely to lead classrooms with higher mean reading fluency estimates, although these findings are not statistically significant among all the subsets of teachers.

The results on state writing assessments were mixed. More time may be required for YRC instructional practices to impact deeper comprehension skills such as inter-textual writing. Anecdotal evidence collected among elementary principals suggests that many teachers are currently focused on reading rate and reading accuracy. It is possible that writing scores have been included in this analysis prematurely, and would be more appropriate as part of a later study.

The reader should note that there is an important segment of data missing from this analysis of the YRC. There are no data related to the depth and quality of classroom implementation of YRC principles. The evaluation design team created an ambitious evaluation design that would analyze the quality of teacher-created lesson plans, instructional practice (including classroom observation by the Tennessee Department of Education Regional office), and the rigor of student work. However, the team was not able to mobilize enough resources to collect representative samples of these data. The evaluation team plans to collect samples of these data in SY1920 using the Achieve.org Instructional Practice Guide (IPG).

References

Abdelrazek, R, (2016). *A Qualitative Analysis of the 2014-2015 Elementary Year Long Reading Course*. Knox County Schools Department of Research Evaluation and Assessment.

Babu, G. J., & Rao, C. R. (2004). *Goodness-of-fit tests when parameters are estimated*. The Indian Journal of Statistics, 66, 63-74.

Binks-Cantrell, E., Washburn, E., Joshi, R., Hougen, M. (2012). Peter effect in the preparation *of reading teachers*. Scientific Studies of Reading, 16(6), 526-536.

Castles, A., Rastle, K., & Nation, K. (2018). *Ending the reading wars: Reading acquisition from novice to expert*. Psychological Science in the Public Interest, 19, 5–51.

Cree, A., Kay, A., & Steward, J. (2012). *The economic and social cost of illiteracy: A snapshot of illiteracy in a global context.* World Literacy Foundation.

Ehri, L. C. (2005). *Learning to read words: Theory, findings, and issues*. Scientific Studies of Reading, 9, 167–188.

Foorman, B. (2016). *Skills to Support Reading for Understanding in Kindergarten through 3rd Grade*. Institute of Education Sciences, U.S. Department of Education, NCEE2016-4008.

Fullan, M. (2007). *Leading in a culture of change*. John Wiley & Sons.

Hall, G. E. (1977). *Measuring Stages of Concern about the Innovation: A Manual for the Use of the SoC Questionnaire*. SEDL.

Hanemann, U. (2015). *Lifelong literacy: Some trends and issues in conceptualizing and operationalizing literacy from a lifelong learning perspective*. International Review of Education, 61(3), 295-326.

Knox County Schools. (2019) *2019-2022 Strategic Plan (Working Copy)*.

Knox County Schools. (2009), *Excellence for All Children: Five Year Strategic Plan 2009-2014.*

Knox County Schools. (2014), *Excellence for Every Child: Five Year Strategic Plan 2014-2019*.

Lyon, G. (2002). *Reading development, reading difficulties, and reading instruction: Educational and public health issues*. Journal of School Psychology, 40, 3–6.

Moats, L. (2009). *Still Wanted: Teachers With Knowledge of Language*. Journal of Learning Disabilities, 42(5), 387–391.

Moats, L. C., & Lyon, G. R. (1996). *Wanted: Teachers with knowledge of language*. Topics in language disorders, 16(2), 73-86.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups*. National Institute of Child Health and Human Development.

Rayner, K., Foorman, B., Perfetti, C., Pesetsky, D., Seidenberg, M. (2001). *How psychological science informs the teaching of reading*. Psychological Science in the Public Interest, 2(2), 31–74.

World Literacy Foundation (2015). *The Economic and Social Cost of Illiteracy*.