Research Study

The Impact of Diagnostic Reviews on Improvement for "Priority" High Schools in Kentucky

October 2022

At the end of the 2011–12 school year, the Kentucky Department of Education (KDE) placed 29 high schools in "Priority" status, defined as schools that had been "Persistently low-achieving". By the end of the 2015-16 school year, 19 of those schools had taken part in Cognia's Diagnostic Review process at some point in the 2012–2013, 2013–2014, 2014–2015, or 2015–2016 school years.[i]

The following report summarizes a quasi-experimental data analysis, showing average increases in Kentucky School Performance Overall Scores for Priority Schools that participated in the Diagnostic Review process.

Diagnostic Review Process

The Cognia Diagnostic Review process provides schools with the support necessary to address underperformance, guide improvement actions, and build leadership capacity in pursuit of meaningful change and effective continuous improvement. This process included three steps implemented at each of the identified Kentucky schools across the 2012 through 2016 school years.

1. During the first step of the Diagnostic Review process, Cognia provided training and support around the eProve™ continuous improvement suite of research-based tools that were used to gather stakeholder input and evaluate evidence of student learning and participation. This training also prepared school leaders with best practices for managing continuous improvement efforts and maximizing results.

2. At the next step, school leaders collected and analyzed qualitative and quantitative evidence in preparation for the onsite Diagnostic Review. Cognia improvement specialists provided coaching on the administration of stakeholder (i.e., parent, student, staff) surveys, completion of a diagnostic self-assessment, and evaluation of student performance data.

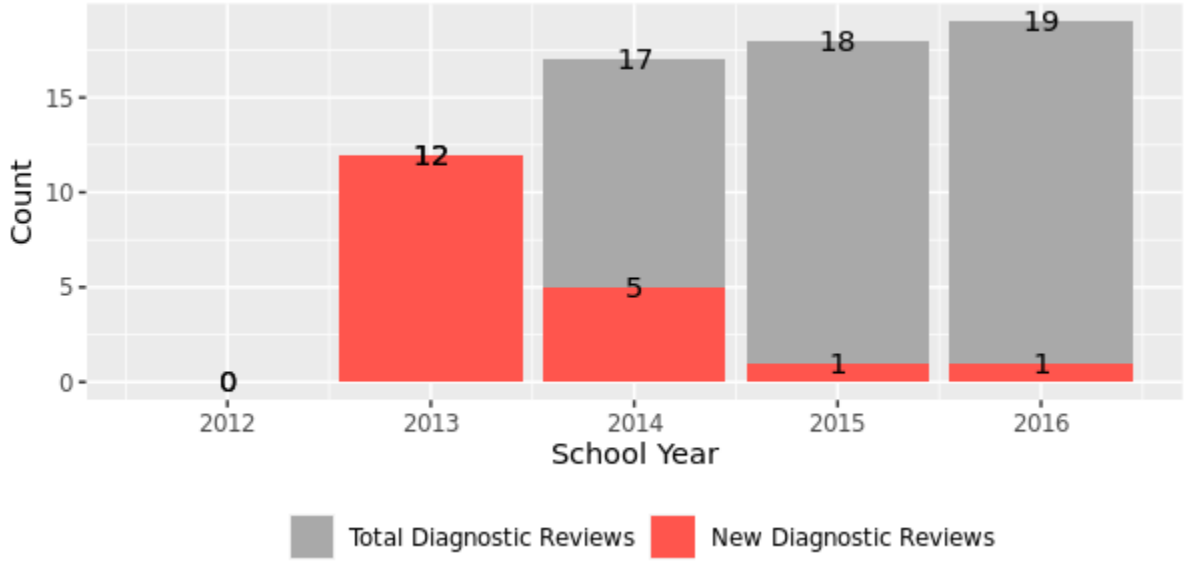
3. The third and final step saw Cognia-certified teams of education experts visiting Kentucky schools to examine the compiled evidence, conduct stakeholder interviews, and observe classrooms using the Effective Learning Environments Observation Tool® (eleot®).

A written report summarizing the findings and conclusions based on the three-day, on-site review was then provided by Cognia following this on-site review. In it, the team identified explicit strengths and challenges along with evidence-based improvement priorities.

Kentucky School Performance

Kentucky Accountability Profiles are publicly accessible and published annually by the Kentucky Department of Education on their School Report Card website[ii]. This data was used to examine whether there were significant differences in gains over time in the Kentucky School Performance Overall Score between "Priority" high schools that participated in the Cognia Diagnostic Review services (19 schools) and those that did not (12 schools).

Beginning with 2012-2013, 12 high schools were identified as Priority and received Diagnostic Review services through to the 2015-2016 school year where one school was identified as Priority and received Diagnostic Review services, bringing the total to 19 schools across the timeline of the state policy. This is a graphical summary of the implementation showing that no Diagnostic Reviews were conducted in the first year of the policy.

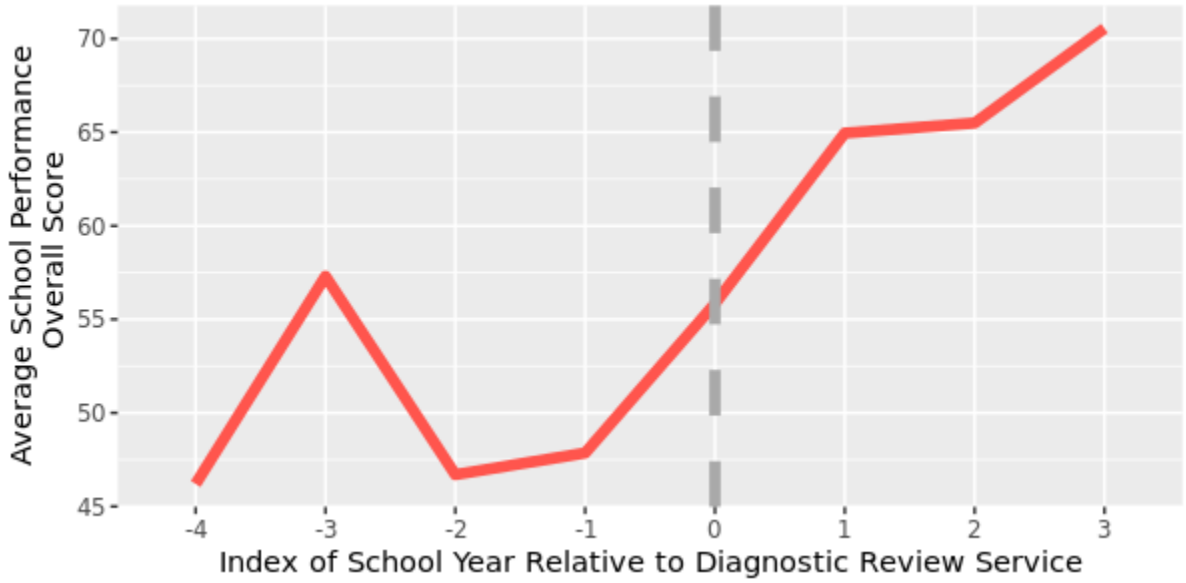


Data Sources

School accountability data, including The Kentucky School Performance Overall Score and demographics, were all drawn from the state department of education website as part of national efforts toward data transparency in accountability. The Kentucky School Performance Overall Score was an index with a value of 0 to 100 reflecting individual schools' academic achievement, achievement gap, growth in reading and mathematics, college/career readiness, and graduation rate.

Demographic information for all schools in Kentucky is similarly available. Before conducting further analysis, the composition of Priority schools was compared against statewide demographic information. A few notable differences between Priority schools and the statewide school characteristics emerge such as Small average school sizes, larger non-white populations and higher rates of Free/Reduced Priced Lunch (see Appendix A).

In the figure below, average School Performance Overall Score by years after delivery of Diagnostic Reviews (i.e., to the right of the dotted line) is greater and demonstrate an increase over average score prior to delivery (i.e., to the left of the dotted line). This simple comparison suggests that Priority schools saw improvement in scores following delivery of Cognia Diagnostic Review services.



Before conducting further statistical analysis it was necessary to demonstrate that there was no change in scores occurring before delivery of the Diagnostic Reviews as this would indicate that score changes could be attributed to external factors. Statistical testing revealed that no significant change was occurring prior to delivery of Diagnostic Review services, therefore, the analysis proceeded.[iii]

Analysis Of Difference-in-Differences

With longitudinal data available and an identification strategy where Priority high schools were assigned to Diagnostic Reviews ("Treatment") on a rolling basis across multiple years, difference-in-differences (DiD) analysis was identified as the appropriate statistical methodology for evaluating impact of Diagnostic Reviews in Kentucky[iv,v]. These 19 schools were assigned to the treatment condition from the point of delivery of Diagnostic Review forward - so, once treated the treatment is maintained.

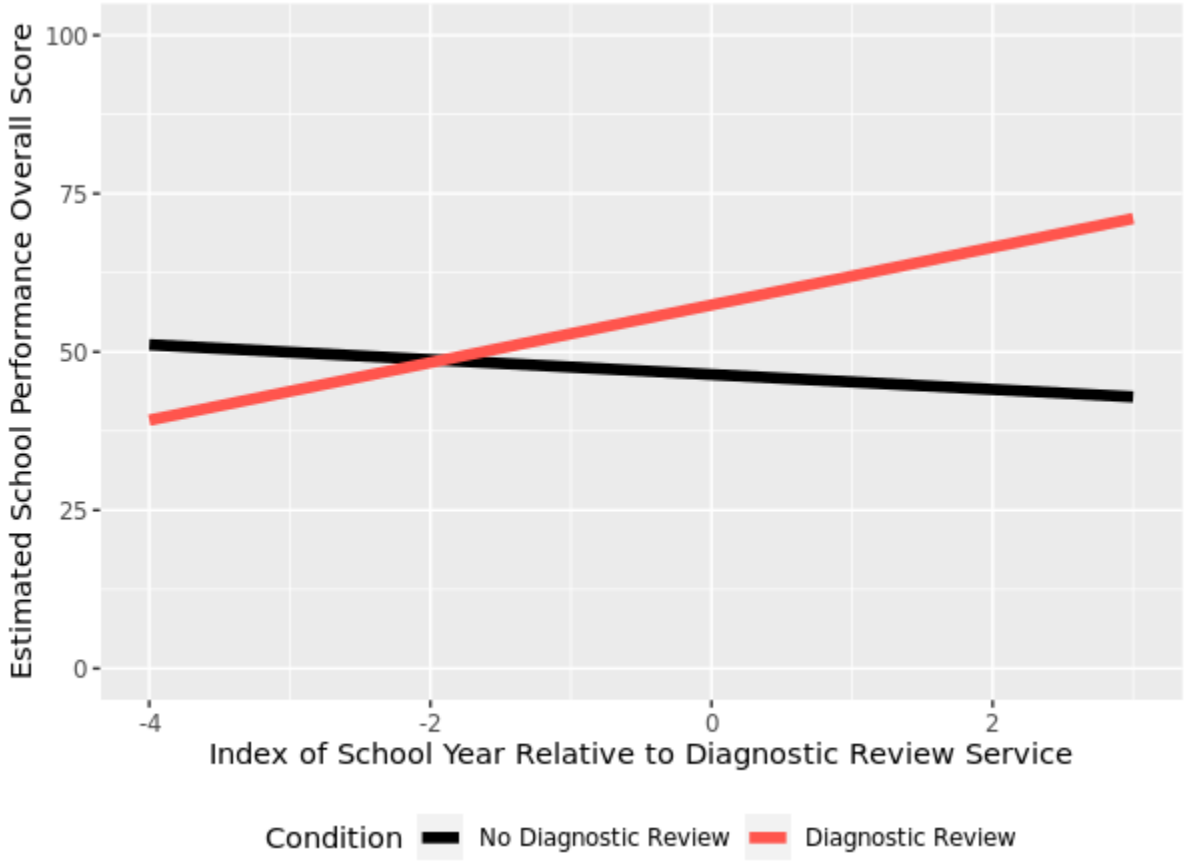
School Performance Overall Scores and demographic values for Treatment schools were submitted to 3 different statistical models to describe the change in scores influenced by Diagnostic Reviews (Appendix B):

- The first model simply examined the effect of treatment as the DiD estimator "D" (Simple Model);
- The second model added the number of school years between each School Performance Overall Score and the first delivery of Diagnostic Review services as an additional factor (Linear Trend Model). These values range -4 to +3, indicating scores before and after treatment; and
- The third model considered the interaction of the treatment and number of school years relative to treatment (Interaction Model).

The final results of all three models show that the treatment effect (DiD) is significant and positive, indicating that Diagnostic Reviews improved the School Performance Overall Scores of treated schools (Appendix C). Of the three models, the Interaction Model (Model 3) was found to fit best, indicating that time relative to Diagnostic Reviews further impacted the scores.

Specifically, schools participating in Diagnostic Reviews experienced gains of nearly 11 points than what were estimated by the statistical model. Further, nearly 5 points additional are gained over the statistical estimates for each year after the Diagnostic Reviews were implemented.

Below is a visual representation of the estimated School Performance Overall Scores from Model 3. As you can see, not only are there notable and significant differences between treatment and control values, but the treatment condition demonstrates a significant gain across years.



Conclusion And Implications

To put these results in context, we look to previous school improvement and turnaround research. Other studies described changes in desired outcomes according to effect sizes, reporting that successful improvement and turnaround efforts demonstrate effect sizes of about 0.100-0.200[vi,vii,viii,ix]; the current study yields an effect size of 0.214 in the first year after implementation. One interpretation of these results is that a school could move from the 50th percentile to the 58th percentile of schools across the state or that a lower performing school could move from the 5th percentile to the 8th percentile, according to School Performance Overall Score.

Federal requirements under the Every Student Succeeds Act (ESSA) require that educational interventions - including turnaround and leadership efforts such as Diagnostic Reviews - demonstrate significant positive impact under robust experimental designs. This study employs such methodology, difference-in-differences, to demonstrate a significant positive effect on School Performance Overall Scores as a result of Diagnostic Review services in the state of Kentucky. Therefore, these results indicate that Diagnostic Reviews are a Tier 2 (i.e., moderate evidence) evidence-based intervention under the provisions of the Every Student Succeeds Act (ESSA).

[i] Note that in 2019 the merger of AdvancED and Measured Progress culminated in the creation of a new organization: Cognia™. Activities described in this paper refer to the organization by its new name.

[ii] https://applications.education.ky.gov/SRC/DataSets.aspx

[iii] This is the assumption of Parallel Trends or Pre-Trend. Satisfying this assumption allows us to causally attribute gains in the accountability index to the treatment - the Diagnostic Reviews. Regression of Accountability Index on years prior to Diagnostic Review produced non-significant results ($\beta = -1.177$, $p = n.s.$).

[iv] Difference-in-differences is an experimental approach that compares changes in an outcome over time when randomization is not possible. Change in the outcome is calculated separately for the treatment and control groups; the change for the control group is then subtracted from the change for the treatment group yielding the "difference-in-differences".

[v] Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. New York, NY: Cambridge University Press.

[vi] Gill, B., Zimmer, R., Christman, J., & Blanc, S. (2007). State Takeover, School Restructuring, Private Management, and Student Achievement in Philadelphia. RAND Corporation.

[vii] Papay, J., & Hannon, M. (2018). The Effects of School Turnaround Strategies in Massachusetts. Presented at the 2018 APPAM Fall Research Conference: Evidence for Action: Encouraging Innovation and Improvement, Appam. Retrieved from https://appam.confex.com/appam/2018/webprogram/Paper26237.html

[viii] Pham, L., Henry, G. T., Zimmer, R., & Kho, A. (2019). School Turnaround in Tennessee: Insights After Six years of Reform. Tennessee Education Research Alliance, Peabody College, Vanderbilt University. https://peabody.vanderbilt.edu/TERA/files/School_Turnaround_After_Six_Years.pdf

[ix] Schueler, B. E., Goodman, J. S., & Deming, D. J. (2017). Can States Take Over and Turn Around School Districts? Evidence From Lawrence, Massachusetts. Educational Evaluation and Policy Analysis, 39(2), 311–332.

Appendix A: Demographic Comparison between all Kentucky Schools and Priority High Schools

	Statewide		Priority	
	N	Percentage	N	Percentage
Total Schools	230		31	
Total Students	141,920		11,578	
Average Enrollment	617.04		373.48	
Male	72,873	51%	6018	52%
Female	68,958	49%	4887	42%
White (Non-Hispanic)	118,118	83%	4090	35%
African American	15,166	11%	1149	10%
Hispanic	4153	3%	117	1%
Asian	1765	1%	0	0%
American Indian or Alaska Native	220	0%	0	0%
Native Hawaiian or Other Pacific Islander	100	0%	0	0%
Two or more races	2398	2%	31	0%
Migrant	131	0%	0	0%
Limited English Proficiency	1472	1%	235	2%
Free/Reduced-Price Meals	72,726	51%	7735	67%
Disability-With IEP (Total)	13,719	10%	1331	11%

Appendix B: Statistical Models to Estimate Difference-in-Differences Effect

[1] Simple model: $y_{jk} = \beta_0 + \beta_1 D_{jk} + \varepsilon_{jk}$

[2] Linear trend model: $y_{jk} = \beta_0 + \beta_1 D_{jk} + \beta_2 T_k + \varepsilon_{jk}$

[3] Interaction model: $y_{jk} = \beta_0 + \beta_1 D_{jk} + \beta_2 T_k + (\beta_3 D_{jk} T_k) + \varepsilon_{jk}$

Where y_{jk} is the School Performance Overall Score for school j at year k, β_0 is the intercept, β_1 is the DiD estimator or treatment effect, D_{jk} indicates delivery of treatment of school j at time k, β_2 is the impact of time T_k, β_3 is the interaction of delivery of treatment and time, and ε_{jk} is the normally distributed error term.

Appendix C: Regression results

	Model 1	Model 2	Model 3**
DiD Estimator	15.315	5.462	10.993
	(2.093)	(3.208)	(3.812)
Time Trend		3.499	-1.177
		(0.904)	(2.051)
DiD*Time			5.727
			(2.270)
N Observations	95	95	95
R-squared	0.365	0.454	0.490

*** Model 3 preferred over Models 1 and 2 (F = 14.979 and 6.366, respectively, df = 1, p < 0.05).*