Supporting Middle School Content Teachers Transition to the Common Core: The Implementation and Effects of LDC

Joan Herman and Scott Epstein

Study Overview and Purpose
States and districts across the country are grappling with the challenge of helping their teachers to transition to the new, more rigorous Common Core State Standards (CCSS) in English Language Arts (ELA) (CEP, 2012). For secondary content teachers now expected to integrate literacy development into their classes, the challenge is daunting indeed. The Literacy Design Collaborative (LDC) provides one possible strategy to help these teachers, and CRESST’s LDC evaluation, the subject here, can provide insights not only about LDC intervention, but the opportunities and obstacles for any transition effort.

LDC supports secondary teachers’ CCSS transition by providing flexible, module templates to enable teachers to seamlessly incorporate ELA into their disciplinary instruction. End-of-module writing tasks provide the heart of the approach. Teachers use fill-in the blank templates to design extended writing tasks that integrate relevant disciplinary content with CCSS literacy demands. An LDC-specified “instructional ladder” helps teachers to design and deliver instructional tasks to provide students the requisite content knowledge and literacy skills to successfully complete the writing task. Core steps of the ladder scaffold student learning and provide ongoing opportunities for formative assessment. Teachers participating in LDC implement at least two modules during the academic year, one targeting explanation, and another targeting argument.

This paper reports on one of two studies CRESST is conducting on LDC. The study reported here examines the implementation and impact of LDC in eighth grade social studies and science classrooms in two states: Kentucky and Pennsylvania. The second study examines district-wide implementation of LDC in sixth grade Advanced Reading classes in a large diverse school district in Florida. Both studies aim to answer the following research questions:

• What is the effect of LDC on student learning?
• How do teachers implement LDC?
• How does the fidelity of LDC implementation affect student learning?
• What conditions and contexts influence LDC learning?

Background
We locate our research in the literature on instructional change. Although classroom practice is notoriously impervious to reform (Cuban 1982; Lortie 1975), an emerging body of research has documented the relationship between student achievement and specific instructional practices that create “opportunities-to-learn” (see Bryk, 2010; Rowen & Correnti 2009; Winters & Herman, 2011). Our implementation focus is classroom instruction, while recognizing that multiple factors influence and inhibit teacher innovation and instructional change (see, for example, Desimone et al. 2002; Fullan, 2008; Owston, 2003). Our ultimate goal is to leverage these findings to inform efforts required for teachers to successfully implement rigorous CCSS demands.
Data and Methodology
The study focuses on multiple measures of both teacher implementation of LDC and student learning. All Kentucky and Pennsylvania eighth grade social studies and science teachers who participated in LDC in 2012-13 and had prior experience with the intervention were invited to participate in the CRESST study. Participation involved sharing information on implementation through teacher logs, surveys, and classroom artifacts, as described below, and administering a CRESST-designed Integrated Learning Assessment (ILA). Participating teachers were asked to focus on one classroom for both the implementation measures and the ILA.

Implementation measures included:

1. *Web-based teacher log.* Study teachers were asked to complete a log twice weekly during their implementation for each of two LDC modules, one implemented in the fall and the second in the spring. The logs focused on (a) the degree to which instruction generally aligned with structure of the LDC intervention, (b) the degree to which instruction explicitly specifies and addresses the discrete literacy skills required to complete the summative writing task, and (c) the quality and extent of formative assessment practices incorporated into LDC instruction. Each log was designed to capture classroom instruction on the particular day the log was completed and focused on only one of the teacher’s classes—the same class for all logs. The log included opening (gateway) items that asked teachers which component of the LDC module they addressed on that particular day (i.e., Preparing for the Task, Reading Process, Transition to Writing, Writing Process) and then branched to “back end” items for the identified component(s), where teachers answered additional questions about component instructional objectives and strategies. For example, the Reading Process section asked teachers to check all the specific reading skills that were addressed in the day’s instruction. A follow-up item asked teachers to identify how they responded if a student had difficulty with the reading assignment. The emphasis on formative assessment aligns with LDC intent, which views the steps in the instructional ladder as opportunities for the teacher to track student progress and intervene appropriately to support student learning. Logs were analyzed at the teacher level.

2. *Analysis of LDC modules.* As part of the log process, teachers in Kentucky and Pennsylvania were asked to upload or otherwise transmit to CRESST the LDC module they had just completed with samples of student work. Modules were scored using a 9-dimension rubric developed by CRESST addressing the quality of the central writing task, the quality of the instructional ladder, and overall quality. Each dimension was scored on a 1-5 scale and both content and literacy rigor were built into the dimensions. Modules were scored by specially trained teachers of the relevant subject area (i.e. social studies/history or science).

3. *Post-intervention Opportunity to Learn surveys.* Participating teachers in Kentucky and Pennsylvania were asked to complete a short one-page survey at the time that they administered the CRESST-designed ILAs. This survey included several questions on the administration of the assessments (i.e., the administration date and amount of time students spent completing them), and a several questions asking teachers to summarize their LDC instruction (number of modules taught in 2012-13 and topics covered). The survey also asked teachers to report on the degree of emphasis they placed in their 2012-13 instruction on the content areas covered by the ILAs: the Reconstruction
period in American History, and scientific theory of evolution. The answers to these questions provide valuable data on students’ Opportunity to Learn (OTL) these content areas, and help to tease out the effects of OTL and LDC instruction.

4. **Teacher surveys.** CRESST collaborated with RFA on the design of the 2013 implementation and scale up survey for teachers. The survey included a section on module implementation with items designed to mirror the intent of the log items. These survey items addressed relative time spent on the various module components, relative emphasis given to specific reading and writing skills, use of formative assessment, strategies for providing feedback, and perceptions of LDC impact. Further, we draw on RFA survey variables as context and possible moderators of LDC implementation and impact—e.g., experience using LDC, attitudes regarding literary instruction, extent of professional development, leadership support, and collaboration. Descriptive statistics were computed at the teacher level.

Measures of student learning include both state assessment scores and performance on CRESST-designed ILAs. Each LDC teacher participating in the study was asked to administer ILAs in just one classroom—the same classroom she logged on. Therefore student results on the ILAs come from a sample of students being exposed to the intervention in target grades and districts, rather than the population of all treated students. The CRESST ILAs are designed to measure both students’ literacy development relative to the CCSS in English Language Arts and depth of content understanding in social studies/history or science and thus are well aligned with LDC goals. Across content areas, the ILAs feature a consistent structure: students first read several texts that typify those encountered in the discipline and address important content; students then respond to reading comprehension and analysis questions that are aligned with the CCSS in ELA. Finally, students respond to an essay prompt that, consistent with the CCSS, asks them to synthesize what they know with what they have read to produce an extended explanation or argument about the topic. Social studies teachers administered an assessment on the Reconstruction period in American History, while science teachers implemented an assessment on the theory of evolution. Essays were scored using specially trained raters and an analytic scoring rubric.

CRESST’s evaluation employs a quasi-experimental design (QED) to compare the performance of LDC students on state assessments to the performance of matched, demographically similar, control students under similar teachers in comparable schools and districts. The pool of treatment students in these analyses include all students taught by LDC social studies and science teachers, not just teachers that chose to participate in the implementation portion of the study. The QED analysis in Kentucky drew comparison students from schools and districts throughout the state, controlling for student demographics, student prior performance, teacher prior effectiveness, and school prior effectiveness. Outcome measures for the Kentucky analysis included eighth grade reading, writing, and social studies state assessment scores. Data for an analysis of the impact of LDC in Pennsylvania is not yet available.

The sample for the Kentucky QED analysis included over 2,200 students taught by 37 teachers implementing LDC in 2012-13, and between 12,000 and 19,000 control students depending on the outcome measure. About half of the 37 teachers completed logs, submitted artifacts, and administered the ILA, while three quarters completed the RFA/CRESST survey. In
Pennsylvania, 22 teachers contributed implementation information, representing over 80 percent of teachers eligible for the study.

Results

Results are promising. Based on descriptive analysis of log, artifact and survey data, fidelity of implementation was quite high in both states. LDC teachers reported in logs and surveys that they addressed a wide variety of reading and writing skills and used a variety of formative assessment strategies in their instruction. Attitudes about the usefulness and effectiveness of the intervention were also very positive. However, there was substantial variation in the specific literacy and formative assessment strategies that individual teachers implemented. Results also showed considerable variation between social studies/history and science teachers, with evidence suggesting that science teachers struggled more with embedding literacy instruction into their content area instruction. Analysis of classroom artifacts revealed that LDC modules on average were moderately well realized, but quality varied greatly across teachers.

Performance on the CRESST ILAs varied considerably across students in both states and assessments. However, overall reading and writing scores on both assessments were quite low. In both states, students earned about half of the possible score points on the ILA reading components. On the writing component, the average student received between 36 and 43 percent of the total possible score depending on the state and assessment.

CRESST used a number of techniques, including factor analysis and cluster analysis, to construct implementation variables from data collected in logs, surveys, and classroom artifacts. Although some interesting patterns emerged, HLM and ANOVA techniques did not find statistically significant relationships between identified implementation variables and student learning. Descriptive results showed that the teachers that scored the lowest on implementation measures also had the lowest scoring students, suggesting there may have been relationships between implementation and student learning that we could not detect due to data limitations, such as sample size.

Results for the impact of LDC treatment on KPREP reading, writing, and social studies are displayed in Tables 1, 2, and 3. Our sample of teachers include all eighth grade social studies and science teachers in five target Kentucky school districts who began teaching LDC in either 2010-11 or 2011-12 and continued implementing LDC in 2012-13. This group included thirty seven teachers, of whom seven were Phase 1 (began LDC participation in 2010-11) and thirty were Phase 2 (began participation in 2011-12). We employed a matching technique known as Coarsened Exact Matching (CEM) to identify comparison students from school districts across Kentucky. Variables included in the matching process included student demographics (gender, race/ethnicity, free/reduced price lunch status, English Language Learner status, etc.), student prior achievement, the effectiveness of teachers prior to the LDC intervention, and school effectiveness. Separate matching processes were employed for each of the three student outcomes.

Subsequent to matching, hierarchical linear models (HLM) were employed to produce estimates of the effect of LDC treatment on student achievement. In the models that produced the results below, unique students were at Level 1 and combinations of social studies and science teachers
were at Level 2. A student received a value of 1 on the treatment status variable if both her social studies and science teachers implemented LDC, 0.5 if one of the two teachers implemented LDC, and 0 if neither teacher implemented LDC. The HLM models controlled for the demographic and achievement variables included in the matching process, making the analysis double robust. In addition, the models included a number of interaction effects between treatment status and demographic or achievement characteristics. These interaction variables explore how treatment effects are modified by student differences on important background variables.

Table 1 displays effect estimates for the impact of LDC treatment on student achievement on the KPREP reading test. Please note that Table 1 (and likewise Tables 2 and 3) only presents results for variables related to treatment status. As noted above, the models also controlled for the main effects of a number of demographic and achievement characteristics. All student achievement variables are measured in standard deviation units based on statewide means. As seen in Table 1, a small positive effect of LDC treatment on KPREP reading test scores was found. The model estimates that a student with LDC teachers in both social studies and science scored .052 standard deviations higher on KPREP reading than a similar student whose social studies and science teachers did not implement LDC. In addition, we found a number of statistically significant interaction effects. The interaction effect with prior student achievement suggests that students that performed at higher levels in the prior year received a greater benefit from the LDC intervention than students on lower levels of the achievement spectrum. Interestingly, there was also a positive interaction effect with free/reduced price lunch status. That is, LDC seemed to have a greater positive impact on students participating in the National School Lunch Program (NSLP) than on students not participating in the NSLP. On the other hand, a moderate negative interaction effect was found for Special Education students, suggesting that Special Education students’ KPREP reading scores did not benefit from LDC. No interaction effect was found with gender.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC Treatment</td>
<td>0.052*</td>
<td>0.024</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Prior Student Achievement</td>
<td>0.032*</td>
<td>0.011</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Free/Reduced Price Lunch Status</td>
<td>0.052*</td>
<td>0.017</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Special Education Status</td>
<td>-0.110*</td>
<td>0.035</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Female</td>
<td>-0.004</td>
<td>0.017</td>
</tr>
</tbody>
</table>

*indicates effect estimate is statistically significant at the 5 percent level.

In contrast to reading, no statistically significant main treatment effect or interaction effects were found for KPREP writing scores, as displayed in Table 2. There is no evidence from our analyses of an impact of LDC on writing performance.
As with writing, no statistically significant main effect of LDC treatment was found for KPREP social studies scores. Our analysis did, however, find interaction effects with prior student achievement and free/reduced price lunch status similar to those found in the KPREP reading analysis. Again, LDC treatment seemed to be more beneficial for students with higher prior achievement and for students participating in the NSLP.

Table 3
Effect Estimates of LDC Treatment in Kentucky Social Studies and Science Classes on Eighth Grade KPREP Social Studies Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDC Treatment</td>
<td>-0.026</td>
<td>0.022</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Prior Student Achievement</td>
<td>0.048*</td>
<td>0.017</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Free/Reduced Price Lunch Status</td>
<td>0.040*</td>
<td>0.020</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Special Education Status</td>
<td>-0.003</td>
<td>0.038</td>
</tr>
<tr>
<td>Interaction: LDC Treatment by Female</td>
<td>0.013</td>
<td>0.016</td>
</tr>
</tbody>
</table>

*indicates effect estimate is statistically significant at the 5 percent level.

Conclusions and Next Steps

In summary, teachers generally had positive attitudes regarding LDC and typically implemented the intervention with high fidelity, although there was variation in focus on literacy skills, use of formative assessment strategies, and quality of modules. QED results should be interpreted with caution given limitations of the study, including teacher sample size and generalizability (results are for one grade and two subjects in one state). Positive QED results in reading are, however, encouraging for an intervention still in its early stages. The positive interaction effect with
free/reduced price lunch status is also encouraging, suggesting that the introduction of literacy instruction into content area classrooms through the LDC intervention may particularly benefit students from low socio-economic backgrounds. Interaction effects with prior achievement and Special Education status suggest that more work may need to be done to tailor the intervention to students with varying levels of prior achievement and special needs. In addition, descriptive results on the CRESST ILAs demonstrate that despite encouraging early results for LDC, students are still a long way from mastering the rigorous literacy skills embedded in the CCSS.

Next steps in the CRESST’s evaluation work on LDC include completing a parallel study evaluating the implementation and impact of LDC in 6th grade English Language Arts classes in a large diverse Florida school district.

References


