Classroom Trials:

A Study of Instruction with Writing Software

January 2016

SRI Education™
A DIVISION OF SRI INTERNATIONAL

Authors

Lauren Cassidy, Kaily Yee, Rebecca Schmidt, Sara Vasquez,
Barbara Means, & Andrew Krumm, **SRI Education**

Developed by SRI Education with funding from the
Bill & Melinda Gates Foundation.

This report is based on research funded by the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

Suggested Citation

Cassidy, L., Yee, K., Schmidt, R., Vasquez, S., Means, B., & Krumm, A. (2016). *Classroom Trials: A Study of Instruction with Writing Software.* Menlo Park, CA: SRI International.

Contents

Executive Summary

The Common Core State Standards (CCSS) in English Language Arts/Literacy (ELA/Literacy) call for changes in literacy instruction to emphasize evidence-based writing and the integration of reading-to-learn, research, and writing. To help their students meet these expectations, teachers will need to integrate reading and writing instruction with the teaching of academic subject matter; they will also need to have their students engage in nonfiction writing more often and at greater length than they have in the past. Doing this will require teachers to read, grade, and provide substantive feedback on more—and longer—student essays.

The burden that these activities impose on middle and high school teachers, who may be grading essays from as many as 150 students, is considerable. As a result, teachers are tempted to give few writing assignments of significant length as a strategy for controlling the amount of time they have to spend grading and providing feedback. Now, however, commercial writing instruction software products have become available that can reduce this burden by organizing the writing process and providing targeted feedback on drafts of student compositions. These products are thought to encourage teachers to assign more writing because they alleviate the associated burden of an initial assessment, enhance student learning by providing immediate feedback on each draft, encourage students to increase revision of their writing, and facilitate managing the writing, feedback, and revision processes. These systems assess the structure and complexity of a student's writing but allow the student to use or ignore the counsel of the software. In either case, as an instrument to aid instruction, the software can produce data that teachers can use to identify the challenges that an individual student or group of students face and to modify their instruction to deal with issues the software has identified. While many of these writing instruction products have been in development for decades, the current national shift toward more extended writing assignments presents an opportunity to investigate these products and their potential to support students and teachers working toward higher performance expectations.

Although such writing software products are being increasingly adopted nationwide, relatively little research has addressed how students and teachers actually use them. With funding from the Bill & Melinda Gates Foundation, SRI Education in collaboration with The Common Pool (TCP) and the Literacy Design Collaborative (LDC) conducted the Classroom Trials study to examine how a sample of middle school teachers implemented writing software in their classrooms. We studied three software products—*Criterion* (ETS), *WriteToLearn* (Pearson Education), and *PEG Writing* (Measurement Incorporated); each product was chosen as a result of an open competition led by TCP. The participating companies applied, and a panel of experts with experience in writing instruction and the assessment of student writing chose the finalists.[1]

The Classroom Trials research was guided by the following research questions:

- How do teachers use writing software to support their writing instruction?
- What challenges do teachers encounter in using these tools?

[1] Additional details about the selection process are available at: https://classroomtrials.rampit.com/#

- What are promising practices for using these tools to support writing instruction?
- How did teachers learn exemplary practices for using these writing software tools?

The Classroom Trials study enlisted two sets of teachers. In the initial phase, the three product vendors nominated 15 experienced teachers who knew how to use their products well. In the second phase, LDC recruited 106 teachers who participated in the Collaborative, which involved teachers in co-design activities and professional learning groups supporting implementation of the CCSS ELA/Literacy.[2] These teachers then participated in a half-day training session, with a separate session for each of the three software products with the expectation that they would implement the product starting in fall 2014. Phase 2 was the main data collection phase.

The Classroom Trials study used the following types of data:

- Teacher interviews and classroom observations
- Daily logs of writing instruction activities
- Online survey of LDC teachers
- Backend system data from the three software products.

Evaluation Findings

In reviewing the Classroom Trials findings, it is important to note that our study focused on identifying key variables that influenced instruction and teachers' perceptions of effects on students across all three writing software products. The writing products were used in different districts, by different teachers, and with different kinds of students. When findings differed across products, it was not possible to determine whether the differences should be attributed to the product per se or to differences in the implementation context.

Characteristics of the Classrooms Taught by LDC Teacher Survey Respondents

The majority of LDC teachers (71%) in the Classroom Trials taught regular, middle school ELA. Overall, these **teachers characterized their students as less proficient in writing than they were in reading.** Nearly half of teachers said they had students reading at or above grade level, whereas **only 36% of teachers reported that their students were writing at or above grade level.**

The grades and student skill levels that LDC teachers taught varied for the three software products. Teachers in the *PEG Writing* sample were more likely than those in the *Criterion* or *WriteToLearn* samples to teach honors and advanced classes, and teachers in the *WriteToLearn* sample had a slightly higher proportion of English Language Learners and a much higher proportion of students writing or reading below grade level in their classes. Not surprisingly, teachers in the *PEG Writing* sample described their students as significantly more proficient in writing than did teachers using the other two products; *WriteToLearn* teachers described their students as significantly less proficient in both reading and writing than did teachers working with the other two software products.

General Writing Instruction

To understand the use of the writing software products, we first gathered data from the teacher logs, surveys, and interviews. That information provided the context needed to ascertain how teachers conducted writing instruction in general, both with and without the writing software. Questions about general writing instruction were used to establish the goals of that instruction, the types of writing that teachers

[2] Although LDC recruited 106 teachers, we invited 100 teachers to complete the survey. Six individuals who were trained in the use of a software product were removed from the sample either because they dropped out of the project or because we learned they were nonteaching staff.

assigned, and the amount of time devoted to writing instruction during a typical week of classes.

The 12 LDC teachers who filled out daily teaching logs reported that they taught writing an average of 2.5 days per week. In terms of elapsed time, LDC teachers logged an average of 87 minutes of writing instruction per week. The time for writing instruction recorded in LDC teachers' logs exceeded the typical time cited in logs completed by the experienced teachers (an average of 55 minutes) participating in phase 1. This difference may be attributable to the LDC emphasis on longer and more challenging writing assignments.

In describing their goals for writing instruction, **LDC teachers' most frequently reported goals were giving students practice with the forms of writing encountered on standardized tests (75%) and improving writing mechanics (73%).** The first of these responses suggests that teachers may have viewed the writing software as attractive because of the similarity of its prompts, readings, and text screens to what students already see on their current writing assessments, e.g. short answers or document-based questions, or will see in the future on new end-of-year accountability tests. The next most commonly reported goals for writing instruction were to explain or analyze a concept or relationship; to reflect on an experience or topic; and to describe a thing, place, or procedure. The goals for writing instruction did not vary significantly by writing software product or course level (honors, regular, or remedial).

To gauge how typical the LDC teachers' goals for writing instruction were, we compared their data with those from two rounds of teacher survey data collected in 2008 and 2010 from much larger samples of teachers participating in the National Writing Project (NWP) (Gallagher, H., et. al., 2009; Gallagher, H., et. al., 2011). Compared with the earlier NWP samples, Classroom Trials LDC teachers put more emphasis on having their students: (1) practice forms of writing they would encounter on standardized tests; (2) explain or analyze a concept (e.g., lab reports); and (3) make a persuasive argument. These differences may reflect recent shifts in writing instruction in response to the CCSS ELA/Literacy's greater emphasis on evidence-based argument writing and the use of writing as part of the process of mastering subject-matter content.

Data from the teacher logs were quite consistent with these teacher survey reports. Teacher logs indicated that short-response, informative, argument or opinion, and journal writing were the most common assignments over the 4-week period that logs were kept. Teachers were less likely to assign creative or narrative writing, which again may reflect the changing CCSS emphasis on nonfiction writing.

LDC teachers indicated that during class time devoted to writing, **students spent most of their time composing text, followed by brainstorming, organizing ideas for writing, and revising text.** LDC teachers said they had their students engage in these activities an average of 2 to 3 times per month; the NWP teachers had reported fewer instances per month.

In terms of length of writing assignments, **the majority of Classroom Trials teachers typically assigned writing of a page or less**—45% said their longest writing assignment was 2 to 3 paragraphs in length, and 26% said it was 1 page. Assignments longer than 1 page were less common, although LDC teachers of more advanced students were more likely than other LDC teachers to report making writing assignments of 2 to 3 pages in length.

Instruction with Writing Software in Classroom Trials Classrooms

On the Classroom Trials survey, **80% of LDC teachers trained on a specific writing software said that they had their students use that software.** From these software users, we collected data concerning the details of their software implementation through additional survey items, daily activity logs, backend user data, interviews, and classroom observations. Topics covered by the data collections were: (1) where and in what configuration they used the software; (2) time spent using the software; (3) their instructional purposes for using writing software; and (4) how they and their students used the software.

An overwhelming majority of these teachers (96%) reported that their students used the software individually, which required a 1:1 computer-to-student ratio. Most *Criterion* and *PEG Writing* teachers used the software in a computer lab or media center equipped with enough computers for all their students. Most *WriteToLearn* teachers, on the other hand, used the software in their regular classrooms, and in half of those cases they did not have enough computers for every student to use the software at the same time. Thus, *Criterion* and *PEG Writing* teachers had to coordinate with other teachers to provide their students with time for using the software in the computer lab, whereas *WriteToLearn* teachers often had to rotate computer use in their classrooms.

In terms of the length of instruction on a given assignment, **most teachers interviewed reported that their students spent a week or less on each writing assignment in which software was used**. The backend user data supported this finding—**across all three products, students spent an average of 3.32 days using the software on each assignment.**

Using the backend user data, we analyzed the average length of each writing assignment composed by students using the software. The average word length across products was 254 words—just over 1 double-spaced page. Given that 71% of teachers responding to the survey said their longest typical assignment was a page or less in length, **the system log data are not inconsistent with the premise that students wrote longer pieces when using the software.**

To investigate how the LDC teachers used the writing software in their writing instruction, we asked them about the software's role in five instructional tasks: (1) prewriting; (2) composing and revising; (3) providing feedback on writing; (4) personalizing instruction; and (5) managing writing assignments. The survey broke each of these instructional tasks down into subtasks, and we combined the teacher responses on the subtasks to produce a single measure for each instructional task.

On average, teachers rated composing and revising, followed by providing feedback, as the most important roles for the writing software. Following these tasks, in order of importance, were personalizing instruction, managing writing, and prewriting. This pattern of responses was similar across the three products, although teachers using *Criterion* reported that the software played a significantly smaller role in prewriting and providing feedback than did *WriteToLearn* and *PEG Writing* teachers. On the survey, *Criterion* teachers reported spending less time in general having their students engage in brainstorming or organizing ideas (with or without writing software). Most *Criterion* teachers interviewed also reported that they tended to have their students use pencil and paper in planning their essays rather than the writing software.

On average, teachers rated the software's role in supporting prewriting tasks as minor. Among those

teachers who did have their students use the software for prewriting activities, **the most prominent prewriting subtasks were outlining and student goal setting/time management.** It is unclear whether the limited use of software in most prewriting subtasks should be attributed to a mismatch between the capabilities of the software and the prewriting activities typically assigned by teachers, to the lack of access to the computer lab, or to the teachers' lack of familiarity with the full capabilities of the software.

More than 75% of teachers said the software played a moderate to major role in supporting students' in all five composing and revising subtasks (composing, editing, revising drafts, mastering basic skills, learning the mechanics of paragraph writing, and providing rubrics and examples to guide students' writing). The backend user data supported this finding: it showed that, on average, **students submitted 3.4 drafts for each assignment,** ranging from 2.20 drafts for students using the *Criterion* software, to 3.89 drafts from *PEG Writing* students.

The majority of teacher survey respondents also said the writing software played a moderate to major role in each of 10 subtasks related to providing feedback on student writing. Nearly all software-using teachers employed the software's machine-scoring feature to provide students with feedback on grammar, mechanics, organization, and sentence fluency. In addition, more than 60% used the software's comment or teacher-student dialog feature to augment this automated feedback with their own comments on the content and organization of the writing. Similarly, **half of the teachers who completed logs said they used the software's "comment" feature to provide students with feedback, and half of the teachers said they also provided in-person feedback to supplement the automated scores generated by the software.** Hence, these **LDC teachers used the software's automated feedback but did not rely on it as the sole mechanism for giving students information about the quality of their writing.** Interviewed teachers said that they were grateful that the software could provide feedback on grammar, usage, and mechanics, thus allowing them to focus on more substantive issues regarding the writing content.

Slightly more than half of the teachers completing the survey reported they used the software to have their students provide feedback to other students. However, analysis of the backend user data on peer reviews from *Criterion* and *PEG Writing* suggested that students using the software generally did not provide in-depth feedback on other students' writing. On average, students using *Criterion* provided less than 1 in-line comment (0.32), and students using *PEG Writing* only completed about 30% of the reviews for their peers that teachers assigned. Moreover, **the peer feedback that students provided was short—just 4.56 words on average.** To enhance support for students in providing more useful feedback to their peers, some of the teachers interviewed scaffolded this skill.

A majority of teachers reported that the software played a moderate to major role in scaffolding and diagnosing differences in abilities or learning gaps, and nearly all teachers said the software played a moderate to major role in letting each student progress at his or her own pace on the same assignments. On the other hand, fewer than half of teachers said that the writing software played a moderate to major role in providing opportunities to incorporate knowledge or interests related to students' backgrounds into their writing or in providing different writing assignments for students with different ability levels.

This finding aligns closely with data from the teacher logs. Teachers who completed the logs reported that they were more likely to use the writing software to provide the same activities for all students (67%) than to provide different activities for different students (33%). Some LDC teachers interviewed noted that using the software to assign different activities to different students was difficult; accordingly they may not have been able to or known how to differentiate activities even if they had wanted to do so.

Teachers most commonly reported that software played a moderate to major role in monitoring student progress over the course of one assignment (88%), managing multiple drafts (79%), and providing scoring supports (79%). The LDC teachers completing the teacher logs most commonly used the software to track student completion of assignments (75%).

Challenges in Using the Software

Survey respondents were asked to rate the extent to which a number of potential challenges affected their use of the writing software. **More than half of all LDC survey respondents said that insufficient class time (60%) and limited student access to technology (56%) moderately to significantly reduced their use of the writing software.** Across all software products, there were few significant differences in the barriers cited.

Other challenges posed a moderate to significant challenge for smaller proportions of teachers. About a fourth, for example, concluded that the software's inability to score: (1) the prompts they developed (27%); and (2) longer writing assignments (24%) were moderate to major challenges. Scoring was of a particular concern for some of the teachers we interviewed. Those teachers wanted their students to use the software to write essays for LDC modules, but the software offered limited feedback on prompts that were not prepopulated in the system. In addition, several teachers said that the scoring was too easy, especially so because the software could not score content; others worried that the scoring was too hard for their less-advanced students.

The key challenges cited by those teachers who had been trained on the software but had not used it with students at the time of the teacher survey were quite similar to those identified by software users. Most often, these nonusers cited the barriers of a lack of time to fit writing software into their curriculum (46%), followed by lack of technology access (38%).

Teachers' Satisfaction with the Writing Software

LDC teachers had favorable impressions of the writing software and its effect on their writing instruction. A large majority of the survey respondents (more than 80%) agreed or strongly agreed that the software improved their ability to:

- Provide more immediate feedback to their students
- Provide individualized supports
- Manage extended writing assignments
- Differentiate content
- Manage student collaboration
- Provide useful writing examples
- Provide useful writing rubrics.

In addition, **more than three out of five teachers in the survey agreed or strongly agreed that the software led them to assign more writing.** Expanding on the survey reports, nearly all teacher interviewed said that they found the instantaneous feedback the software provided to students on spelling, grammar, and other writing conventions—an aspect of grading that is often time-consuming—to be one of the most valuable features of the software

programs because it gave teachers time to focus on other aspects of student writing.

LDC teachers were generally positive about the effect of the writing products on student outcomes. **Nearly all surveyed teachers agreed or strongly agreed that the software gave students more opportunities to edit their work (96%), improved student engagement (94%), encouraged peer-to-peer feedback (92%), improved students' writing (92%), and gave students more opportunities to practice writing (91%).** Several teachers interviewed noted differences in the impact of these products on students of varying achievement levels. Some teachers cited particular advantages for lower performing students, noting that the software built up their confidence in their writing by allowing them to see improvements in their scores immediately. Both LDC and experienced **teachers across all three products cited the appeal of the "gaming" aspect of these products, which motivated even reluctant writers to revise their essays.** Finally, **two-thirds of teachers responding to the survey said they were "very likely" to recommend the writing product to a colleague.**

Teacher Recommendations for Additional Support and Program Improvements

LDC teachers were trained to use the writing software primarily during a half-day training session in the fall of 2014. The amount and content of this training were similar to those that the training vendors normally provide for new customers. LDC teachers rarely received follow-up training and support from product vendors, LDC, or their districts and schools, however. Instead, most LDC teachers relied on their colleagues, when they were available, for help or used a trial-and-error approach to learn more about the nuances of the software and determine how to incorporate writing software into their instruction.

Teachers generally reported that they were comfortable with learning how to implement the writing software through trial and error with informal help from colleagues. Several teachers did note, however, that **they would have liked more training about using: (1) the writing software in settings with limited computer access; (2) specific features of the software (e.g., peer review); and (3) the formative data that resulted from use of these systems.**

Research on teacher learning has demonstrated the value of school-based communities of learners focused on implementing particular changes in teacher practice (Bryk et al., 2010; National Academies of Science, Engineering, and Medicine, 2015). The Classroom Trials study design and recruiting process did not support the development of such communities at participating schools. In a grade-level or schoolwide adoption of a writing software product, opportunities for teachers to support each other and exchange implementation tips and learning activities might well enhance usage rates and student outcomes.

Apart from the training, teachers provided the following suggestions for **improving the software by enhancing flexibility to adapt its products and its features to meet students' unique needs:**

- Making the software's prewriting tools more adaptable for teacher-developed assignments
- Adding to software capability to give students different reading and writing assignments depending on their ability
- Providing the option to change the feedback automatically provided to students in order to tailor feedback better
- Enabling students to view their prewriting work on a split screen while they were composing to allow easy reference to both documents at the same time.

Compatibility of the Writing Software with the CCSS and the LDC Approach to Literacy

For teachers implementing curriculum modules developed through LDC, the writing software was typically used at the end of the process when students had to write essays. Teacher interviews suggested several areas where the compatibility of the writing software with the LDC approach to literacy instruction could be improved. **LDC emphasizes evidence-based writing and connecting writing to research, and a number of teachers felt the software products could do more to support students in that kind of work.** Teachers wanted to insert their own prompts and reading materials easily and to have the system provide the full range of rubric scores for writing to these prompts (a capability beyond some existing systems, which must be "trained" using a set of student writing samples). They also wanted to allow students to view their prewriting products and their unfolding composition at the same time so that they could move back and forth between the two. Teachers also wanted more support for the conventions of evidence-based writing, including proper citation practice. Teachers also noted that the products often could not verify whether citations were done properly or treated them as grammar errors.

Promising Practices in Incorporating Software into Writing Instruction

A major goal of the Classroom Trials was to identify promising practices for use of writing software in support of CCSS in ELA/Literacy. Practices that SRI researchers judged to be worthy of emulation and further research included:

- Using software-generated rubric scores to identify common writing weaknesses that can be addressed in teacher-led lessons before students continue work on their essays
- Using system data to group students with similar skill needs for small-group work on their common weakness
- Having students complete a reading, followed by group discussion of the reading content, and then writing in response to a prompt using the software
- Modeling informative feedback on writing so that students can determine what to look for and how to provide feedback that actually informs revisions
- Developing a collection of starter sentences that can scaffold students' efforts to frame constructive feedback on their classmates' writing
- Double-blind peer feedback, with students exchanging their writing and scoring each other's work without knowing their identities
- Using the software to create a portfolio of each student's writing and associated rubric scores over time to illustrate individual student progress and areas for further work.

As these practices in the Classroom Trials illustrate, teachers participating in this study treated the writing software as an amplifier of their own teaching strategies rather than a replacement for their judgment and skills as teachers.

Introduction

The Common Core State Standards (CCSS) in English Language Arts and Literacy (ELA/Literacy) call for changes in instruction to emphasize evidence-based writing and the integration of reading-to-learn, research, and writing. Helping their students meet these expectations will require teachers to integrate reading and writing instruction with the teaching of academic subject matter and to ask their students to engage more often in writing nonfiction passages that are longer than those required in the past. Doing this will require teachers to read, grade, and provide substantive feedback on more—and longer—student essays.

The burden that these activities impose on middle and high school teachers, who may be grading essays from as many as 150 students, is considerable. As a result, teachers are tempted to give few writing assignments of significant length as a strategy for controlling the amount of time they have to spend grading and providing feedback. But there are now commercially available writing instruction software products that could alleviate some of this burden by organizing the writing process and providing targeted feedback on drafts of student compositions. Promotion of these products is based on the belief that they encourage teachers to assign more writing because they alleviate the associated burden of grading, enhance student learning by providing immediate feedback on each draft, encourage students to do more revision of their writing in order to improve their scores, and facilitate managing the writing, feedback, and revision processes.

These systems assess the structure and complexity of a student's writing but allow the student to use or ignore the counsel of the software. In either case, as an instrument to aid instruction, the software can produce data that teachers can use to identify the challenges that an individual student or group of students face and to modify their instruction to deal with issues the software has identified.

Although many of these writing software products have been in development for decades and are being adopted at growing rates across the country, relatively little research has addressed how students and teachers actually use them. Some studies have tested the accuracy of machine scoring of students' writing, compared with scoring by trained humans (Preston & Goodman, 2012; Shermis & Hamner, 2013). A study of the use of *Criterion* in the 2002-03 school year (Attali, 2004) found that 71% of student essays produced within the software were submitted for scoring only once, undermining the rationale that software feedback could improve student writing. Warshauer and Grimes (2008) studied the use of two other writing instruction software systems, and similarly found that 72% of essays were submitted for scoring only once. They also found a low level of software usage in the three middle/junior high schools and one high school participating in their study: on average classes used the system for less than a half essay per month.

Neither of these two studies examined how particular features of writing software products are being used in combination with activities outside the software in different phases of the writing process. The Classroom Trials study was designed to investigate how teachers manage the writing process, provide feedback to students both within and outside the software, and use software-generated data to help target their instruction. The national shift towards

more writing assignments presents an opportunity to investigate these products and their potential to support students and teachers working toward higher performance expectations.

Purpose of the Study

SRI Education, in collaboration with The Common Pool (TCP) and Literacy Design Collaborative (LDC), received a grant from the Bill & Melinda Gates Foundation to study how middle school teachers use writing software in their classrooms. This Classroom Trials study was designed to inform school administrators, teachers, and students seeking to implement the CCSS in ELA/Literacy by highlighting productive models for writing instruction that capitalize both on the capabilities of the writing software and on the practices of experienced users of the software and reform-oriented teachers.

Specifically, we studied the use of three products—*Criterion* (ETS), *WriteToLearn* (Pearson Education), and *PEG Writing* (Measurement Incorporated)—which were chosen through an open competition led by TCP. The participating companies applied, and a panel of experts with experience in writing instruction and the assessment of student writing chose the finalists.[3]

The Classroom Trials research was guided by the following research questions:

- How do teachers use writing software to support their writing instruction?
- What challenges do teachers encounter in using these tools?
- What are promising practices for using these tools to support writing instruction?
- How did teachers learn exemplary practices for using these writing software tools?

The Classroom Trials study consisted of two sets of teachers. In the initial phase, the three product vendors nominated 15 experienced teachers who were well-versed in using their products. For the second phase of the study, LDC recruited teachers who participated in the Collaborative, which involved teachers in collaborative activities and professional learning groups supporting implementation of the CCSS ELA/Literacy. For each of the three software products these teachers participated in a separate half-day training session, with the expectation that they would implement the product starting in fall 2014.

We hoped that the experienced teachers' participation would serve to identify promising practices of teachers who were comfortable using the products to support writing instruction. From the LDC teachers, we expected to learn how the software products could best be used in classrooms where teachers focus on addressing the CCSS in ELA/Literacy. By engaging these two teacher groups, the Classroom Trials study aimed to describe use models applicable to a wide range of classrooms across the country and to examine the specific functionalities of the writing software products that support writing.

This final report (1) describes how teachers used the software products in writing instruction, (2) identifies challenges to using the software, (3) summarizes teacher recommendations for additional support and software improvements, and (4) describes promising practices for incorporating software into writing instruction.

[3] Additional details about the selection process are available at: https://classroomtrials.rampit.com/#

Characteristics of Good Writing Instruction

To determine how writing software can be used in the classroom and those features of writing instruction that it is best designed to support, the characteristics of writing instruction that have been found to be effective in improving students' writing skills need to be addressed.

The literature on writing instruction identifies seven elements of good practice. These elements are described by Graham (2008) and confirmed in other studies (Graham, S., Gillespie, A., & McKeown, D., 2013; Graham & Perin, 2007; Graham & Sandmel, 2011; Zumbrunn & Krause, 2012). The elements follow:

Providing Ample Time for Writing. To improve their nonfiction writing, students need to spend time writing in all of their core subject area classes. In fact, research on the development of writing proficiency suggests that students should be writing daily for multiple purposes, including communication, persuasion, reflection, mastering of content material, and demonstration of knowledge. Writing activities should increase in complexity as students advance through the grade levels (Graham, 2008; Graham et al., 2013; Zumbrunn & Krause, 2012).

Enhancing Students' Knowledge of Writing. Students can learn about the elements of good writing and the purposes and styles of writing in multiple ways, including reading good writing; discussing the characteristics of strong writing; imitating effective writing; and reading to gain content knowledge regarding the topic of their writing (Graham & Perin, 2007).

Encouraging Students' Engagement in Writing. Students need to be interested in and enjoy writing to be motivated to improve their writing. Effective ways of increasing student engagement include allowing students to choose writing topics of interest to them, providing purposeful assignments, allowing students to collaborate on writing, and setting high but attainable expectations for students' writing (Zumbrunn & Krause, 2012).

Scaffolding Students' Work Throughout the Writing Process. Students need support as they learn to work through the multiple components of the writing process. It is not enough for teachers to simply assign students to prewrite, draft, revise, edit, and share their writing. The teachers should model and teach strategies that students should use in each stage of the writing process and then reduce external supports as students gradually gain independence in executing the processes (Graham, 2008).

Perfecting Students' Basic Writing Skills. Graham (2008) argues that students need to master basic skills, such as spelling and punctuation, to free mental resources to address other areas of competent writing, such as voice, fluency, and word choice.

Using Technology. In addition to technologies designed to support writing for students with special needs, word processors can facilitate writing for all students by making revisions easy, allowing students to produce professional looking papers, and lessening the burden on students with fine motor challenges (Graham, 2008; Graham & Perin, 2007).

Enabling Learning through Timely Feedback. People learn better when they receive immediate feedback (Bransford, Brown, & Cocking, 2000). Using feedback, teachers and students alike can modify their practice and provide timely assessments of student writing. Students can benefit from teacher, peer, and self-feedback on their writing, especially in relation to specified writing goals that they can then work to improve (Graham, 2008; Zumbrunn & Krause, 2012). For teachers, the information gained by assessing their students' writing allows them to refine their teaching to meet their students' needs better.

Characteristics of Good Writing

Good writing instruction and assessment require an understanding of what constitutes good writing for a specific purpose and audience. While recognizing the differences in writing for different contexts, writing instructors have worked to define dimensions of quality that apply to all writing. A common framework for analyzing the quality of student writing is the 6+1 Trait® Writing (ideas, organization, voice, word choice, sentence fluency, conventions, and presentations) created by Education Northwest curriculum experts in collaboration with teachers in the early 1980s.[4]

Common Core State Standards in English Language Arts and Literacy

Released in 2010, CCSS in ELA/Literacy have shifted the focus of writing instruction in many schools. Compared with most prior state standards, the CCSS for ELA/Literacy place more emphasis on reading complex informational texts, using academic language, and relying on evidence from the texts to support analyses and claims. An assumption underlying the CCSS is that tasks and texts should grow in complexity across the grades and build students' knowledge across the content areas.[5]

EngageNY succinctly summarized six key shifts in writing instruction called for by the CCSS for English Language Arts, as shown in Exhibit 1.[6]

The three vendors whose products were included in Classroom Trials have taken the CCSS into consideration as they continue to develop and refine their writing software products.

Exhibit 1. Shifts in ELA/Literacy Instruction Called for in the CCSS for ELA/Literacy

Shifts in ELA/Literacy		
Shift 1	Balancing Informational and Literary Text	Students read a true balance of informational and literary texts.
Shift 2	Knowledge in the Disciplines	Students build knowledge about the world (domains/content areas) through text rather than the teacher or activities
Shift 3	Staircase of Complexity	Students read the central, grade-appropriate text around which instruction is centered. Teachers are patient, create more time and space and support in the curriculum for close reading.
Shift 4	Text-based Answers	Students engage in rich and rigorous evidence-based conversations about text.
Shift 5	Writing from Sources	Writing emphasizes use of evidence from sources to inform or make an argument.
Shift 6	Academic Vocabulary	Students constantly build the transferable vocabulary they need to access grade-level complex texts. This can be done effectively by spiraling like content in increasingly complex texts.

[4] See http://educationnorthwest.org/traits/about-61-trait-writing for more information.

[5] See http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/ for more information.

[6] This chart was retrieved from https://www.engageny.org/resource/common-core-shifts.

The Classroom Trials Writing Software Products

Writing software products for inclusion in the Classroom Trials were chosen through an open competition led by TCP, a management consulting firm that specializes in helping clients design incentive models to attract diverse participation in efforts to solve challenging issues. TCP began the preparatory work for these Classroom Trials in June 2012, with support from the Carnegie Corporation of New York and the William and Flora Hewlett Foundation. Over the course of 14 months, TCP solicited applications from every known software provider supporting K-12 writing instruction, focusing specifically on software systems capable of delivering automated feedback and assessment of writing artifacts generated by students in a classroom setting. To guide this process, TCP formed a working group, which included representatives with expertise in automated scoring (e.g., credible academics familiar with those vendors and their track records), teachers (e.g., from both the National Education Association and American Federation of Teachers), and others. The Working Group identified three characteristics as most important when attempting to support writing instruction (classroom compatibility, student engagement, and instructional support), which became the basis for scoring the software products in the competition.

Through this process, three product providers were selected to participate in the Classroom Trials: ETS (*Criterion*), Measurement Incorporated (*PEG Writing*), and Pearson Education (*WriteToLearn*). Each product is briefly described next, along with a description of ideal product use, according to each software provider's representatives. Exhibit 2 summarizes the features of each product.

Criterion

Criterion is a Web-based learning tool designed to support writing instruction. Once teachers have created their assignments using either prompts preloaded into *Criterion*, which are aligned with the CCSS ELA/Literacy standards, or their own prompts, students can enlist a range of planning tools before writing their essays. Once a piece of writing is submitted, the software generates a holistic score based on an analysis of student performance on the widely used six traits of writing (ideas, organization, voice, word choice, sentence fluency, and conventions), as well as a count of the number of errors on these traits, with feedback on each error, including suggestions for revision. In addition to the system-generated feedback, teachers can add their own comments, to which students can respond using the program, engaging the teacher in a dialog about the essay and the feedback.

According to the ETS representatives SRI interviewed, teachers should use *Criterion* as a supplemental tool that allows students to practice and improve their writing. It is not intended to replace the teacher, and each teacher should come up with her own model for how the software should be used in her classroom.

PEG Writing

PEG Writing is a writing practice website developed by Measurement Incorporated to help students in Grades 3 through 12 improve their writing skills and inform instruction. To assign essays to students, teachers either select from *PEG Writing's* prepackaged writing prompts or create their own. Students can then use the software's prewriting tools to plan their essays, and when a student submits his

essay, *PEG Writing* scores the essay on the six traits of writing. Immediately after submitting the essay, the student receives a score report, including a holistic score, trait scores, and grammar and spelling suggestions. With this feedback, *PEG Writing* provides students with links to writing lessons associated with specific skills. Teachers can also provide feedback in a draft and ask other students to submit feedback on a peer's work.

The *PEG Writing* representative interviewed argued strongly that writing needs to be a part of the ELA/Literacy classroom on a daily basis and that it should also be incorporated into all content areas at least monthly. The goal would be to develop common writing prompts for use across the curriculum so that teachers in multiple subject areas can work together to develop good writing. Finally, she noted that students can use the software to practice and strengthen their skills to become better writers.

WriteToLearn

Distributed by Pearson Education, *WriteToLearn* is a fully automated online literacy tool. In addition to building writing skills, *WriteToLearn* is also designed to develop reading comprehension for students in grades 4-12. Using *WriteToLearn*, students complete essays on prepackaged writing prompts relevant to subjects across the curriculum and aligned with the CCSS or with teacher-created prompts. In either case, students receive immediate feedback (holistic scores) on their writing, as well as feedback on the six traits of writing. Teachers can also provide feedback to students directly via *WriteToLearn*.

To develop reading comprehension and writing skills in tandem, *WriteToLearn* enables students to read passages on a variety of topics assigned by teachers and then summarize these passages in their own words. *WriteToLearn* then measures the student's content coverage and provides instantaneous feedback on whether the summary covers the key information of the passage, has been condensed from the original text, and used the students' own words. The software also identifies spelling and grammar errors. Within these reading passages, students also can complete exercises to develop their vocabulary skills.

Representatives of *WriteToLearn* described the ideal way to use the software entailing assigning both summaries of reading passages and essays multiple times a week. *WriteToLearn* representatives acknowledged that although doing so could be challenging because of competing academic priorities, students using the software frequently would benefit from sufficient opportunities to practice writing to improve their skills.

Exhibit 2. Writing Software Features by Product

	Criterion (ETS)	*PEG Writing* (Measurement Incorporated)	*WriteToLearn* (Pearson Education)
Prewriting Tools	Yes	Yes	No
Prepopulated Writing Prompts	Yes	Yes	Yes
Prepopulated Writing Prompts Aligned to CCSS	Yes	Yes	Yes
Teacher-created Writing Prompts Allowed	Yes	Yes	Yes
Tutorials	Not Available	Yes	Not Available
Teacher-to-Student Feedback	Yes	Yes	Yes
Student-to-Teacher Communication	Yes	Yes	Not Available
Peer Review Tools	Yes	Yes	Not Available
Additional Features	Not Available	Not Available	Summary Writing Vocabulary Exercises

Methods

Research Questions

The Classroom Trials research was guided by the following research questions:

- How do teachers use the writing software to support their writing instruction?
- What challenges do teachers encounter in using these tools?
- What are promising practices for using these tools to support writing instruction?
- How do teachers learn exemplary practices?

Sample Selection

This study used convenience sampling to select two distinct groups of writing teachers for participation in the study. In the first phase of the project, software vendors nominated teachers who had been using their three writing software products for some time and were considered proficient and skilled users. We refer to this small sample as "experienced teachers." The second group of teachers was drawn from LDC participants, who have been working to improve implementation of the CCSS for ELA/Literacy and who had not used any of the writing software products before our study. Throughout the report, we refer to this larger sample of reform-oriented teachers who were new users of the three products as "LDC teachers." Please see Appendix A for additional information about these two samples of teachers.

Experienced Teachers

We collected data from experienced users of the products to help pilot survey and observation instruments, raise additional implementation questions for the study, and, to a lesser extent, provide a point of comparison with the sample of LDC teachers. We asked each product vendor to nominate 4-8 teachers who, to the best of their knowledge, represented high-quality use of their writing software. The vendors based their nominations on reports from their sales representatives, who had the most contact with users, as well as information reported by schools and districts. Some of the selected teachers had served as models for effective practice in the past (e.g., they were featured in training materials).

We piloted our data collection instruments with 2 teachers in the winter of 2014. The remaining 13 teachers were asked to participate in a spring 2014 site visit and to complete a record of their writing instruction ("instructional log"). We also analyzed software backend user data for these teachers' classes.

LDC Teachers

Our study focused primarily on the use of the writing software by LDC teachers because of their intense focus on writing as part of a larger effort to teach to the CCSS ELA/Literacy. As teachers who work with LDC, they engage in practices the CCSS encourages, such as assigning longer, evidence-based writing tasks, and integrating writing across the content areas. Aiming to make literacy instruction the foundation of the core subjects, LDC encourages teachers to build content knowledge on top of a coherent approach to literacy.

LDC works with teachers to build students' literacy skills through a teacher-created instructional design system. The LDC Framework consists of modules, which are 2- to 4-week instructional units made up of a teaching task, standards, minitasks, and other instructional elements. Teachers are expected to use this common framework to individually or collaboratively develop

literacy-rich curricula in their content areas. Using the LDC Framework, teachers can merge CCSS literacy standards with subject area standards.

For the main Classroom Trials data collection, LDC recruited 106 writing teachers in districts in three geographic regions already participating in the study, with one of the writing software products offered for implementation in each region. Teachers in these districts were invited to participate in fall 2014 in a half-day training session on the assigned writing software product in their region. The study design called for training 40 middle school teachers per product at each of three sessions. Recruiting proved difficult in the Pennsylvania region, but a small number of high school teachers was accepted to enlarge that region's sample; even so, the final available sample for *Criterion* fell short of the target, as shown in Exhibit 3.[7]

Exhibit 3. LDC District Assignments, Demographics, and Study Participants

Product	District	Number of Teachers Trained	Demographics*		
			% ELL	% SPED	% FRPL
Criterion (ETS)	School District A, Pennsylvania	26	>1%	17%	51%
	School District B, Pennsylvania		1%	16%	24%
	School District C, Pennsylvania		3%	24%	68%
	School District D, Pennsylvania		5%	15%	28%
	School District E, Pennsylvania,		3%	16%	32%
	School District F, Pennsylvania		17%	18%	82%
	School District G, Pennsylvania		12%	20%	84%
	School District H, Pennsylvania		3%	12%	24%
	School District I, Pennsylvania		1%	18%	28%
PEG Writing (Measurement Incorporated)	School District J, Florida	40	12%	14%	57%
WriteToLearn (Pearson Education)	School District K, Louisiana	40	10%	11%	78%
	School District L, Louisiana		5%	7%	67%
	School District M, Louisiana		2%	10%	79%

*Demographic data was retrieved on December 18, 2015 from:
http://nces.ed.gov/ccd/districtsearch/index.asp
http://www.education.pa.gov/Data-and-Statistics/Pages/Loan-Cancellation,-Low-Income.aspx#.VnQyst-rTNQ
http://www.ed.gov/labor-management-collaboration/conference/hillsborough-county-public-schools
http://jpschools.org/about-us/facts-figures/
http://www.coweninstitute.com/wp-content/uploads/2013/07/2013_SPENO_Final2.pdf
http://www.louisianabelieves.com/resources/about-us/10-years-after-hurricane-katrina
http://www.louisianabelieves.com/data/reportcards/2015/

[7] Pseudonyms were used in place of district names to preserve confidentiality.

One hundred LDC teachers who attended these training sessions were asked to complete a survey in December 2014 and January 2015.[8] SRI then used survey results to select 4-8 teachers from each region for visits at their schools (including an interview and a classroom observation) and to complete an instructional log. Teachers were selected on the basis of their reported frequency of use of the product since the training. In all regions, most teachers were trained in September[9], and site visit selection occurred in February or early March 2015, giving teachers at least 3 to 4 months to use the product before they were asked to participate in interviews and observations.[10]

Frequency of use varied across the three software products, but on average, teachers using *Criterion* and *WriteToLearn* were selected if they reported using the product more than twice since training, and teachers using *PEG Writing* were selected if they used the product more than five times total or more than three times per week. Our final site visit and log sample included 7 teachers using *Criterion* (6 middle and 1 high school), 7 using *PEG Writing* (all middle school), and 4 using *WriteToLearn* (all middle school), for a total of 18 teachers. We also reviewed backend user data for all teachers who used one of the products.

[8] Although LDC recruited 106 teachers, we invited 100 teachers to complete the survey. Six individuals who were trained in the use of a software product were removed from the sample either because they dropped out of the project or because we learned they were nonteaching staff.

[9] Due to recruitment difficulties, some *Criterion* teachers were not recruited and trained until November 2014.

[10] In the district using *PEG Writing*, the use of the software was delayed by a month because of problems with student rosters.

Data Collection Activities

To examine the use of writing software to support writing instruction, the study used both qualitative and quantitative data, drawing on background research on the three products, visits to both experienced teachers' and LDC teachers' classrooms to conduct interviews and observations, instructional logs collected from experienced and LDC teachers, surveys of LDC teachers after their first semester of product use, and backend user data obtained from each of the three products. Appendix A describes these data collection activities in more detail.

Site Visits. Members of the SRI research team conducted site visits with 15 experienced teachers in winter/spring 2014 and 18 LDC teachers in spring 2015.

During the visits, we interviewed teachers to learn about how and why they used the writing software products, the supports they found helpful, and the challenges they encountered in learning and using the software. Wherever possible, we also observed the teacher using the product for one class period to connect observed practices with teachers' descriptions of their intents, beliefs, and perceptions, and to understand the contextual factors that influence software use.

Teacher Log. All teachers in the site visit sample were asked to complete an online daily teaching log of their writing instruction to provide a detailed and accurate description of use. We used log data in conjunction with the backend user data (described below) from the same period to help us better understand teachers' purposes for using the software and their rationale for the frequency and pattern of use observed.

Teacher Survey. We administered a survey to LDC teachers to assess how they used the writing software to support their writing instruction, the features of the

software they and their students used, and supports they received to use the tools. We also asked whether they expected to use the writing software in the future. We had an overall response rate of 70%, with higher response rates for the *Criterion* and *PEG Writing* samples. Although the sample size was adequate to capture large differences in how the three products were used, it was not large enough to capture moderate or small differences among products.

Backend User Data Analysis. We collected and analyzed backend system data for the three products to learn how students actually used the tools (e.g. how frequently they visited the tool, the features of the tool they accessed, the amount of time they spent using each software feature).

We completed two waves of backend data analysis. Using data from the classrooms of experienced teachers, we conducted exploratory data downloads and analyses in spring 2014 to familiarize SRI analysts with the data elements available from each writing product and to refine our processes for the second wave of analysis. The second wave of analysis using data from all trained LDC teachers who used the product occurred in spring 2015. Using these data in conjunction with teacher survey responses, we explored patterns in students' use of each system in relation to the instructional goals and practices of their teachers.

All of these data collection activities informed our findings. In cases where different methods provided data on the same variable (e.g., some items in the teacher survey, teacher log, and backend user data), we compared responses across data collection methods as a validity check. In other cases, the analysis of one type of data helped explain findings for another. For example, interviews provided context and insight into some of the survey findings. In the remainder of this report, we present findings from the LDC sample of new users. As appropriate, we draw on data from the smaller sample of experienced teachers to provide context.

Evaluation Findings

In reviewing the Classroom Trials findings, it is important to note that our study focused on identifying key variables that influenced instruction and teachers' perceptions of effects on students across all three writing software products. The writing products were used in different districts, by different teachers, and with different kinds of students. When findings differed across products, it was not possible to determine whether the differences should be attributed to the product per se or to differences in the implementation context.

We begin by describing the classrooms of LDC teachers and their writing instruction in general. Next we examine LDC teachers' use of the software products in their writing instruction, the challenges they encountered in using the software, their satisfaction with the products, and their recommendations for additional support and program improvements. Throughout the report, we highlight examples of software use that we regard as promising, given their consistency with research-informed recommendations for writing instruction.

Characteristics of the Classrooms Taught by LDC Teacher Survey Respondents

Across all three sites, 71% of teachers in the LDC sample taught regular, middle school ELA, as shown in Exhibit 4. Overall, survey respondents characterized their students as more proficient in reading than in writing. Nearly half of teachers said they had students reading at or above grade level, whereas only 36% of teachers reported this level of proficiency in writing (this difference is statistically significant at $p<0.01$).

Exhibit 4. Grade Levels and Characteristics of Classes Taught by Surveyed LDC Teachers

	Percent of Respondents			
	Total Sample	*Criterion*	*PEG Writing*	*WriteToLearn*
Grade				
Sixth	54%	30%	63%	67%
Seventh	43%	40%	38%	56%
Eighth	40%	30%	47%	39%
High School	7%	20%	0%	6%
Subject				
ELA	71%	65%	78%	67%
Mathematics	6%	10%	0%	11%
Science	11%	20%	0%	22%
Social Studies	14%	25%	3%	22%
Other	23%	5%	38%	17%

	Percent of Respondents			
	Total Sample	*Criterion*	*PEG Writing*	*WriteToLearn*
Class Level				
Regular	77%	90%	69%	78%
Honors	53%	35%	72%	39%
Advanced Placement (AP)	9%	5%	13%	6%
Remedial	31%	40%	31%	22%
Other	13%	5%	19%	11%
Student Characteristics				
Special Education	16%	15%	16%	16%
English Language Learners	9%	6%	10%	12%
Proficient in Reading	49%	60%	65%	11%
Proficient in Writing	36%	25%	61%	6%

The grades and student skill levels that LDC teachers taught varied across the three sites (and thus across the three software products). The *Criterion* sample included 3 high school teachers, and the largest group reported that they taught seventh grade (40%). In contrast, the majority of respondents in the *PEG Writing* (63%) and *WriteToLearn* (67%) samples taught sixth grade.

Teachers in the *PEG Writing* sample were more likely than those in the *Criterion* or *WriteToLearn* samples to teach honors classes ($p<0.01$), and teachers in the *WriteToLearn* sample had a slightly higher proportion of English Language Learners and a much higher proportion of students writing or reading below grade level in their classes ($p<0.05$). Not surprisingly, teachers in the PEG Writing sample described their students as significantly more proficient in writing than did teachers using the other two products ($p<0.05$); *WriteToLearn* teachers described their students as significantly less proficient in both reading and writing compared to teachers working with the other two software products ($p<0.001$).

General Writing Instruction

To provide the context necessary to understand the way in which writing software products were used in teachers' writing instruction, we first gathered data from the teacher logs, surveys, and interviews on how teachers conducted writing instruction in general, both with and without the writing software. Questions about general writing instruction were used to establish the goals of that instruction, the types of writing teachers assigned, and the amount of time devoted to writing instruction during a typical week of classes.

Frequency of Writing Instruction

The 12 LDC teachers who filled out daily teaching logs reported that they taught writing an average of 2.5 days per week, with more writing instruction in the first week (3.1 days) than in the fourth week (1.8 days) of log completion. In terms of elapsed time, LDC teachers logged an average of 87 minutes of writing instruction per week. Teachers reported spending more time in writing instruction in week 1 (120 minutes) than in week 4 (62 minutes).

One reason for the difference reported in the amount of writing time may be teachers' desire to comply with expectations for writing instruction in the first week of filling logs and a return to more typical instruction in subsequent weeks. Another reason may have been the date of each state's standardized testing window, which was scheduled for a few weeks after the end of the log period.

On average, the time devoted to writing recorded in LDC teachers' logs exceeded the time recorded by experienced teachers participating in phase 1 (87 minutes for LDC teachers compared with 55 minutes for experienced teachers). This difference may be attributable to the LDC emphasis on longer and more challenging writing assignments.

Writing Goals

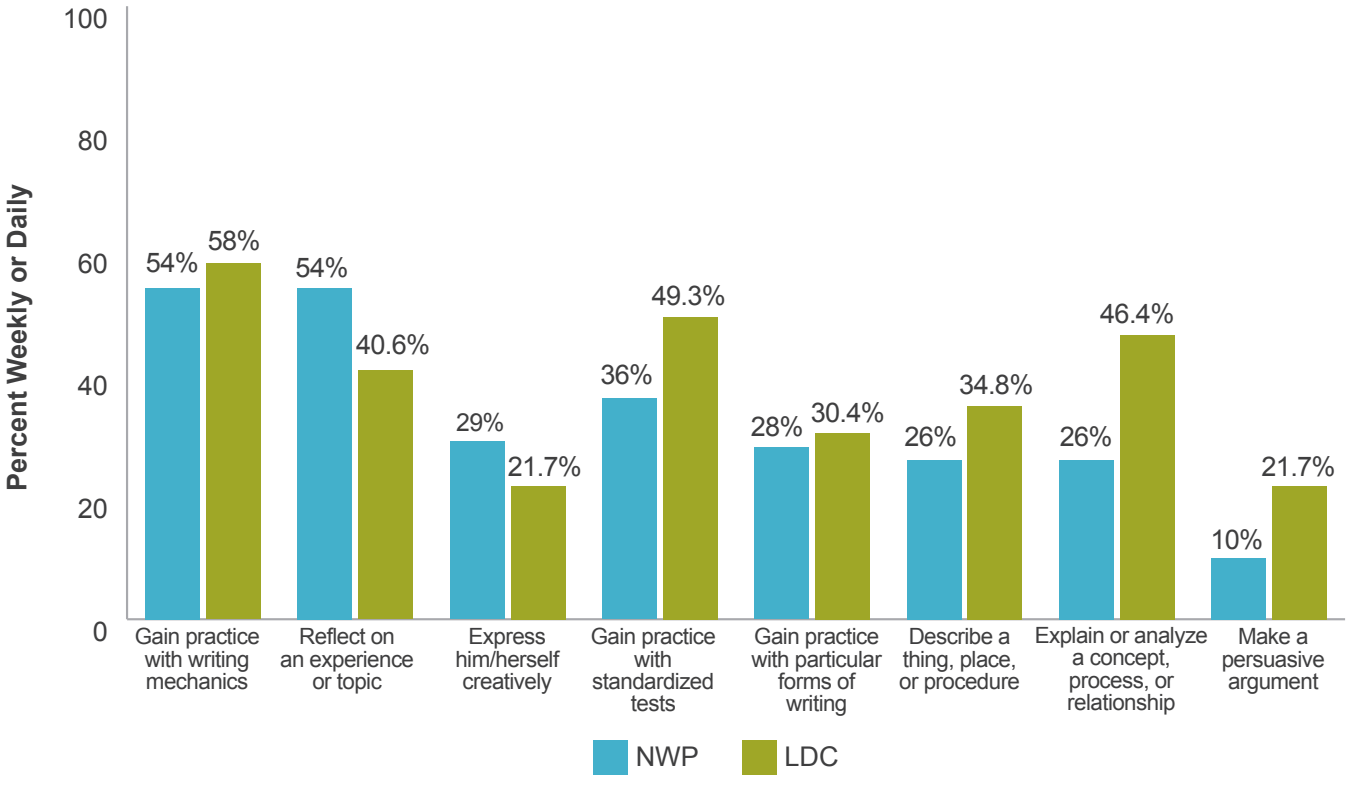
LDC teachers' most frequently reported goals for writing instruction were improving students' writing mechanics (58%) and giving students practice with the forms of writing encountered on standardized tests (49%). The second goal suggests that teachers may have viewed the writing software as attractive because of the similarity of its prompts, readings, and text screens to what students already see on their current writing assessments, e.g. short answers or document-based questions, or will see in the future on new end-of-year accountability tests. The next most commonly reported goals for writing instruction were to explain or analyze a concept or relationship; to reflect on an experience or topic; and to describe a thing, place, or procedure. The goals for writing instruction did not vary significantly by writing software product or course level (honors, regular, or remedial).

Teachers who taught eighth-grade students indicated that they had students write with a higher average frequency across all purposes compared with teachers of other grades ($p<0.05$), perhaps in preparation for the increased amount of writing that students will be assigned in high school.

To gauge how LDC teachers' goals for writing instruction agree with those of other teachers, we compared their responses with those from two rounds of teacher surveys collected in 2008 and 2010 from much larger samples of teachers participating in the National Writing Project (NWP) (Gallagher, H., et. al., 2009). Compared with the earlier NWP samples, Classroom Trials LDC teachers put more emphasis on giving their students practice with forms of writing they would encounter on standardized tests, explaining or analyzing a concept (e.g., lab reports), and making a persuasive argument (Exhibit 5).[11] These differences may reflect recent shifts in writing instruction in response to the CCSS ELA/Literacy's greater emphasis on evidence-based argument writing and using writing as part of the process of mastering subject-matter content.

Data from the teacher logs were quite consistent with these teacher survey reports. Logs indicated that short-response, informative, argument or opinion, and journal writing were the most common assignments over the 4-week period that logs were kept. Teachers were less likely to assign creative or narrative writing, which again may reflect the changing emphasis of writing instruction called for by the CCSS in ELA/Literacy.

Exhibit 5: Writing Goals for LDC and NWP Teachers



[11] We were unable to test these differences for statistical significance because standard errors were not reported in the National Writing Project reports.

Types of writing

LDC teachers indicated that students spent most of the time they devoted to writing to composing text, followed by brainstorming, organizing ideas for writing, and revising text. LDC teachers said they had their students engage in these activities an average of 2 to 3 times per month, as shown in Exhibit 6, compared with an average of just once a month for NWP teachers (Gallagher, H., et. al., 2011).[12] Again, these differences may be attributable to the shift in focus toward writing in the CCSS for ELA/ Literacy since the NWP study was conducted.

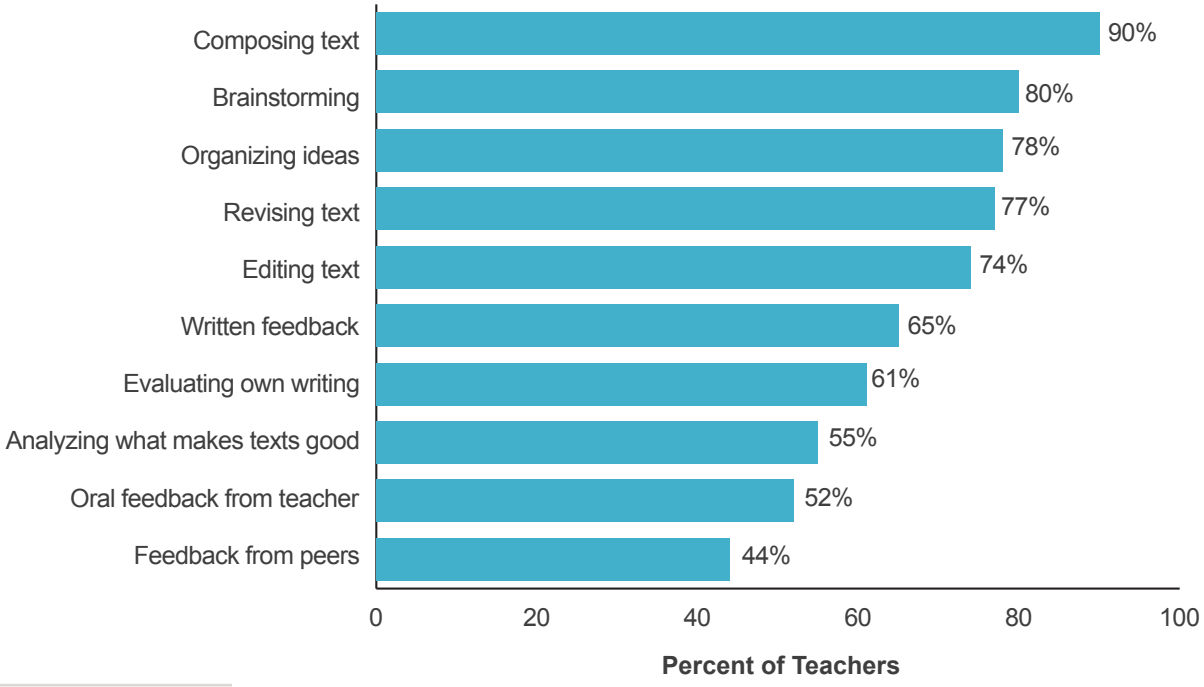
Teachers of eighth-grade students reported a higher average frequency of writing activities during class time overall compared with teachers of other grades ($p<0.05$). In terms of product differences, teachers in the *Criterion* sample were significantly less likely than teachers using the other two products to indicate their students engaged in brainstorming, organizing ideas, or evaluating their own writing ($p<0.05$).

Length of writing assignments

The majority of Classroom Trials teachers typically assigned writing of a page or less. LDC teacher survey responses indicated that the longest writing assigned by these teachers was typically 2 to 3 paragraphs in length (45%), followed by 1 page (26%). Teachers rarely assigned writing of 4 pages or more (10%). Experienced users of the writing software reported writing assignments of similar length.

LDC teachers of honors, advanced, and AP courses were more likely to assign writing of 2 to 3 pages in length than were teachers in regular, remedial, or other courses ($p<0.05$). Teachers in the *PEG Writing* sample, who were more likely to teach honors and advanced students and to report that their students were proficient in writing, were also more likely to assign writing of 2 to 3 pages in length rather than shorter pieces ($p<0.05$). There were no other significant differences.

Exhibit 6: Writing Activities Occurring 2-3 Days per Month or More Often



[12] We were unable to test these differences for statistical significance because standard errors were not reported in the National Writing Project reports.

Instruction with Writing Software in Classroom Trials Classrooms

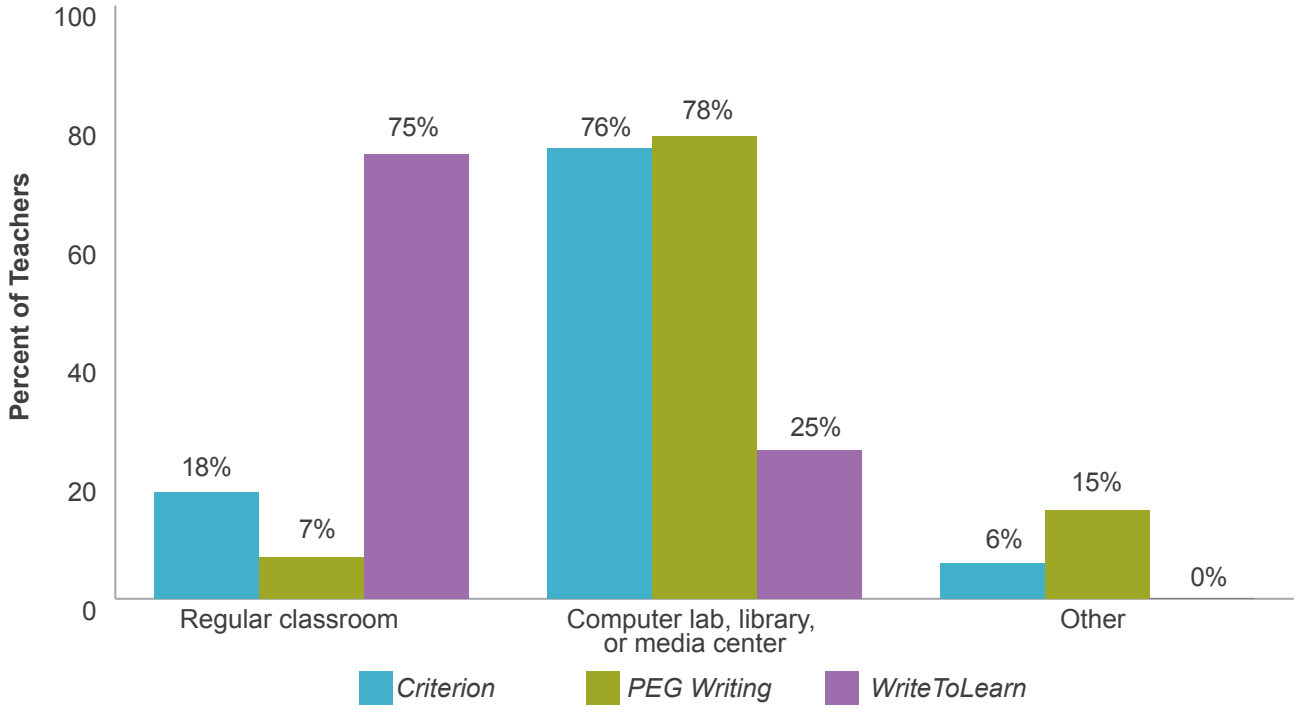
On the Classroom Trials survey, 80% of LDC teachers said that they had their students use the writing software for which they received training. They included 85% and 84% of teachers who volunteered to use *Criterion* (17) and *PEG Writing* (27), respectively, but only 67% of those who signed up to use *WriteToLearn* (12). From these software users, we collected data concerning the details of their software implementation through additional survey items, daily activity logs, backend user data, interviews, and classroom observations. Topics covered by the data collections were: (1) where and in what configuration they used the software; (2) time spent using the software; (3) their instructional purposes for using writing software; and (4) how they and their students used the software.

Location of Software Use

An overwhelming majority of these teachers who used the writing software reported that their students used them individually (96%), which requires a 1:1 computer-to-student ratio. As shown in Exhibit 7, most *Criterion* and *PEG Writing* teachers used the software in a computer lab or media center, which were equipped with enough computers for all their students. Among those who used the software outside their classroom, the ratio of students to computers during software use was 1-to-1 for all *Criterion* users and 96% for *PEG Writing* users.

Most *WriteToLearn* teachers, on the other hand, used the software in their regular classrooms, and in half of those cases they did not have enough computers for every student to use the software at the same time. Thus, *Criterion* and *PEG Writing* teachers had to coordinate computer lab access with other teachers to provide their students with time using the software, whereas *WriteToLearn* teachers often had to figure out a computer rotation schedule within their classrooms.

Exhibit 7: Locations Where Students Typically Used the Software



Length of Instruction per Writing Assignment

Most teachers interviewed reported having their students complete each writing assignment in which software was used in a week or less time. Because many teachers had to rely on using writing software in computer labs, libraries, and other common spaces where they had to share computers with other teachers, most teachers reported that students completed each assignment from prewriting, through editing, to final submission over the course of a week or less. The backend user data collected by all three software systems supported this finding. Across all three products, students spent an average of 3.32 days using the software on each assignment.

The length of time spent on each assignment varied by grade level and by software product. Sixth-grade teachers reported spending the most time on each assignment, logging in to the software on an average of 3.89 days, with seventh and eighth grade teachers reporting slightly fewer days, on average (3.38 and 3.14 days each). Grade 6 had an average of 2.58 days, and high school had the fewest days per assignment, averaging just 1.73. Across products, students using *Criterion* logged the fewest days per assignment (2.67 on average), while students using the *WriteToLearn* software logged the most, at 3.57. *PEG Writing* students logged an average of 3.46 days per assignment using the product. This difference may be attributable to several different factors. For example, *Criterion*'s sample included the most high school teachers, who reported the fewest days per assignment, which may have resulted in a lower average for *Criterion* users overall. Teachers using *WriteToLearn* have more freedom to use computers for a longer period of time as they more often reported using the software on computers in their own classroom. *WriteToLearn* teachers also tended to teach students that were less proficient in reading and writing than teachers using *Criterion* and *PEG Writing*, and so teachers may have given students more time to complete their assignments.

Length of Student Writing with the Software

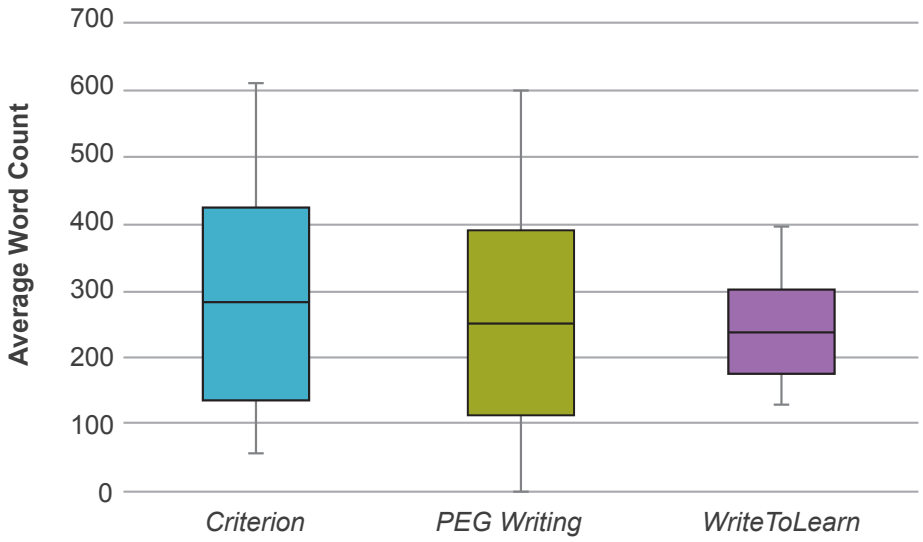
As indicated earlier, LDC teachers most often reported that the longest writing task they assigned in a typical week, with or without the writing software, was 2 to 3 paragraphs in length (45%), followed by one page (26%). Using the backend user data collected from each software system, we were able to analyze the average length of each writing assignment composed by students using the software. The average word length across products was 254 words—just over one double-spaced page. Given that 71% of teachers responding to the survey said their longest typical assignment was a page or less in length, **these system log data are not inconsistent with the premise that students write longer pieces with the software.**

As Exhibit 8 shows, although the length of essays ranged widely, students using *Criterion* and *PEG Writing* produced essays that were, on average, 282 and 252 words long, respectively—just over one double-spaced page in length. Students using *WriteToLearn* had the lowest average word count with 240 words. As noted above, *WriteToLearn* teachers had a slightly higher proportion of English Language Learners and a significantly higher proportion of students writing or reading below grade level. Therefore, writing was likely to be challenging for those students, resulting in shorter essays than students using the other two products.

It should be noted also that the software products have limits on the number of words that can be submitted to satisfy an assignment[13]; the LDC teachers cited the products' limitation on the length

[13] For *Criterion*, the word limit for holistic scoring at the Grade 4-12 level is 800 words. Essays of more than 800 words can be submitted, but those essays will not be scored. *WriteToLearn* does not provide scores for responses longer than 1,200 words. *PEG Writing* does not have a word limit and will score longer essays. *PEG Writing* software evaluates essays by comparing them to essays that have been previously scored by human reviewers. As of summer 2014, new models were implemented based on essays of up to 4,500 words.

Exhibit 8: Average Word Count by Product



of student work as one needing improvement, as discussed later in this report. Nonetheless, the average length of student writing on the software in no case approached the systems' word limits.

Instructional Tasks Supported by Writing Software

To investigate how the LDC teachers used the writing software in their writing instruction, we asked teachers about the software's role in supporting five instructional tasks: (1) prewriting; (2) composing and revising; (3) providing feedback on writing; (4) personalizing instruction; and (5) managing writing assignments. On the survey, each of these instructional tasks was broken down into subtasks, and teacher responses on the subtasks were then combined to produce a single measure for each instructional task (Exhibit 9).

On average, teachers rated the writing software's role highest in composing and revising, followed by providing feedback (Exhibit 10). Overall, the median rating of the role of the software in writing process tasks (composing, editing, basic skills, rubrics, and paragraph writing mechanics) was 3.4 (out of 4), which is between a "moderate" and "major role" across these activities, as shown in Exhibit 10. The medians for feedback (3.0), personalizing learning (2.8), and managing writing (2.8) show that teachers thought the software played a "moderate" role in these activities. On average, teachers rated the software's role in supporting prewriting tasks as minor, with a median rating of 2.1 for prewriting activities. It is interesting to note that 67% of LDC teachers completing a teacher log reported that students used the software to plan at least once over a 4-week period.[14]

As seen in Exhibit 11, the general pattern of responses to the roles that the software played in instruction was similar across the three software products, although teachers using *Criterion* reported that the software played a significantly smaller role in prewriting and in providing feedback on writing than did teachers using the other two products ($p<0.05$). It should be remembered that when responding to

[14] This finding may be attributed to several factors. For example, teachers selected to complete the logs were more frequent users of the software products than were all teachers completing the survey. Several months also elapsed between the time teachers completed the logs and the subsequent survey. During that time, teachers may have become more comfortable using the various features of the writing software product, including the planning/prewriting activities.

Exhibit 9. Instructional Tasks Supported by Writing Software

Instructional Task	Component Sub-tasks of Each Instructional Task
Prewriting	Provide reading comprehension practice Provide vocabulary practice Support student note-taking (e.g., with annotation tools) Answer and/or generate questions about a writing topic Support outlining, planning, and organizing Support student goal-setting and time management for the writing process Promote dialog or facilitate class discussion on a topic
Writing process	Support students' composing of initial drafts Support students' editing and revisions of drafts Provide opportunities for students to master basic writing skills Provide rubrics, models, and examples to guide student writing Support students' skills in the mechanics of paragraph writing
Feedback on writing	Provide automated feedback to students on writing mechanics Provide automated feedback on overall quality of writing Allow students to provide feedback to other students Provide feedback specifically linked to the CCSS Supplement the feedback provided by the software using the comment feature Supplement the feedback provided by the software using the teacher-student dialog Supplement the feedback provided by the software outside of the computer Allow students to self-assess their own writing against specific criteria Use additional prompts to practice areas of weakness identified by the software Set rules for the number of revisions a student may make to a writing assignment
Personalized learning	Provide differentiated assignments for different student abilities Provide practice for some students while the rest of the class is engaged in other activities Provide student choice in activities/assignments Incorporate student background and prior knowledge in learning activities Provide different levels of support or scaffolding so that all students can participate at their level Provide opportunities for students to progress at their own pace Diagnose students' learning abilities and gaps Communicate students' writing progress to parents
Managing writing	Distribute assignments and materials to the class Differentiate assignments on the basis of student ability Help manage the collection and review of multiple drafts and revision Monitor student progress within an assignment Monitor student progress over the course of the semester or year Manage group writing projects Provide supports for the scoring of writing assignments (e.g., providing rubrics) Distribute grades for writing assignments Publish student writing projects to class or others in the school

questions about their writing practices in general, survey respondents using *Criterion* reported spending less time having their students engage in brainstorming or organizing ideas (with or without writing software). Most of the *Criterion* teachers we interviewed also reported that they generally had students use pencil and paper in planning their essays before starting to write their essays on the computer. Two of the *Criterion* teachers noted that they tended not to use the software's prewriting tools because they were not adaptable to their specific teacher-developed assignments and because students could not open multiple windows to view their prewriting and their essay at the same time.

Exhibit 10: Roles of the Software in Supporting Writing Instruction Tasks

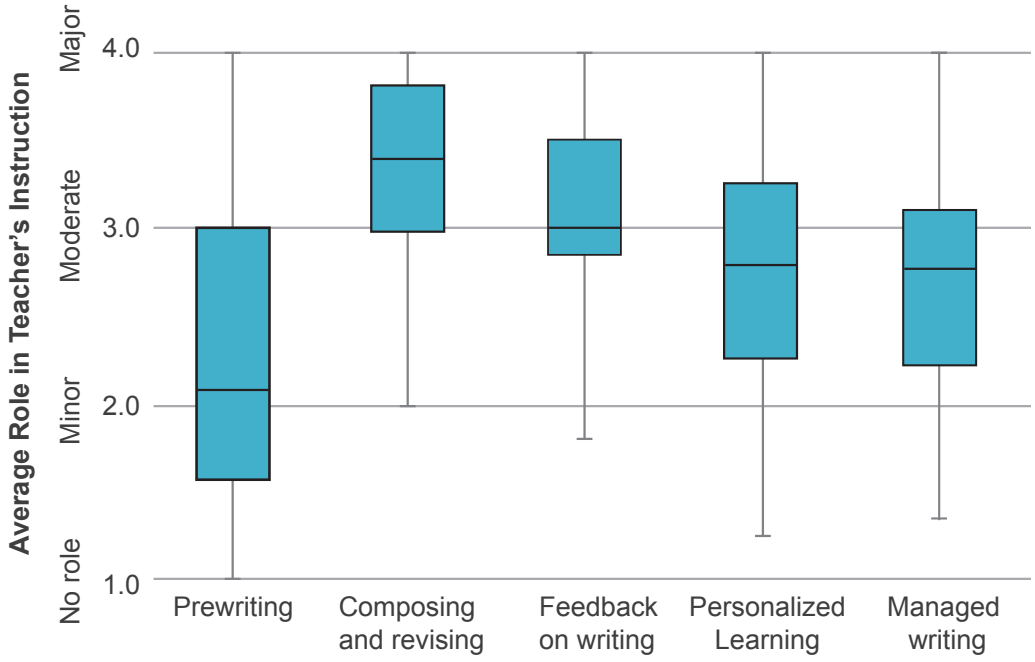
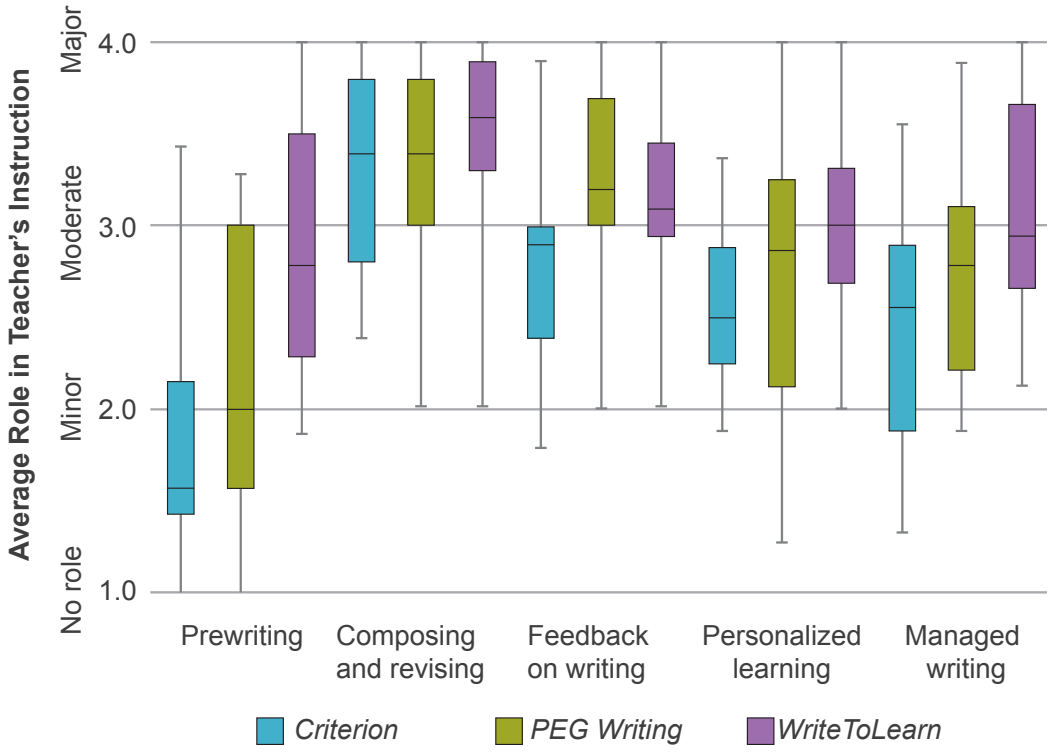


Exhibit 11: Roles of the Software in Supporting Writing Instruction by Software Type



Teachers who used *WriteToLearn* reported that the software played a significantly larger role in prewriting than did other Classroom Trials teachers ($p<0.05$)—a surprising finding, given that *WriteToLearn* does not have an integrated suite of prewriting tools. Printable graphic organizers are available in *WriteToLearn* to help students organize their writing, as are essay tips to provide guidance as students begin their writing; teachers may have considered these two functionalities "prewriting," even though the vendor does not think of its product as supporting this phase of writing. *WriteToLearn* teachers may also have considered close reading of articles pulled from *WriteToLearn* as a type of "prewriting" activity. Students in one classroom we observed, for example, first annotated two articles pulled from *WriteToLearn* by comparing and contrasting the two types of writing described in the articles before they started to use the software to write their essays.

The median ratings of the role of the software in instruction did not differ significantly between teachers who said the average writing proficiency of their students was at or above grade level and those who said it was below grade level. During the site visits, however, teachers did note some challenges with using the writing software with lower achieving students. One teacher using *Criterion* said that the rubric that the software used to evaluate student writing was very challenging for lower level students and that these students tended to score on the lower end of the range, no matter what they did. Other teachers noted that the number of corrections that struggling writers received was often overwhelming and could be discouraging. One *PEG Writing* teacher said that she tended to shut down the software's feedback feature until students had a chance to edit earlier drafts on their own.

Use of Software to Teach Prewriting

As noted, teachers reported that writing software played only a minor role in teaching prewriting. Among teachers who had their students use the software for prewriting activities, the most frequently reported activities were to support outlining (55%) and student goal setting and time management (54%). As shown in Exhibit 12, fewer than half of teachers said the software played a moderate to major role in the other pre-writing tasks addressed on the survey.

It is unclear whether the limited use of software in most prewriting subtasks should be attributed to a mismatch between the capabilities of the software and the prewriting activities typically assigned by teachers, to the lack of access to the computer lab, or to the teachers' lack of familiarity with the full capabilities of the software. (It should be noted that *WriteToLearn* was not explicitly designed to support students' use of the software to complete prewriting activities.)

Instead of using tools for prewriting activities embedded in the software, many of the teachers we interviewed reported that they tended to have students use pencil and paper for prewriting. For example, one of the *WriteToLearn* teachers we observed directed students to create a Venn diagram on paper to help them compare and contrast what they read in two articles about zoos' approaches to caring for their animals. Once students completed the diagrams, they used them to start drafting their essays using the writing software. Often this tactic of prewriting using pencil and paper was used because of limited access to the computer lab, but in one district teachers reported that they were more likely to use a district-created prewriting template with their students, and other teachers said that their students preferred completing their prewriting on paper rather than the software.

Exhibit 12: Teachers' Ratings of the Role of the Software in Supporting Prewriting as Moderate to Major

Use of Software to Teach Composing and Revising

More than 75% of teachers said the software played a moderate to major role in all five composing and revising subtasks (composing, editing, revising drafts, mastering basic skills, mastering the mechanics of paragraph writing, and providing rubrics and examples to guide students' writing), as shown in Exhibit 13. In addition, **nearly 90% of teachers said that they had their students use the software to support editing and revising**. Teacher ratings of the software's importance for these process activities did not vary significantly by software product.

Both the teacher log and backend user data supported this finding. A significant majority of the LDC teachers who completed the teacher logs reported that students used software when composing text (67%) and revising text (83%) at least once during the 4 weeks the logs were maintained. The backend user data showed that **students submitted 3.4 drafts for each assignment, on average,** ranging from 2.20 drafts completed by students using the *Criterion* software to 3.89 drafts from *PEG Writing* students (Exhibit 14). This finding stands in sharp contrast to those of earlier studies in which most students submitted their writing for scoring only once and did no online revision (Attali, 2004; Warschauer & Grimes, 2008). Most of the teachers we interviewed said that they did not limit the number of times students could revise their writing and submit it for feedback. On the other hand, at least two teachers using *PEG Writing* set a *minimum* number of drafts that each student had to submit per assignment in order to make their students revise their drafts at least once and encourage them to submit their drafts for peer review.

Exhibit 13: Teachers' Ratings of the Role of the Software in Supporting the Writing Process as Moderate to Major

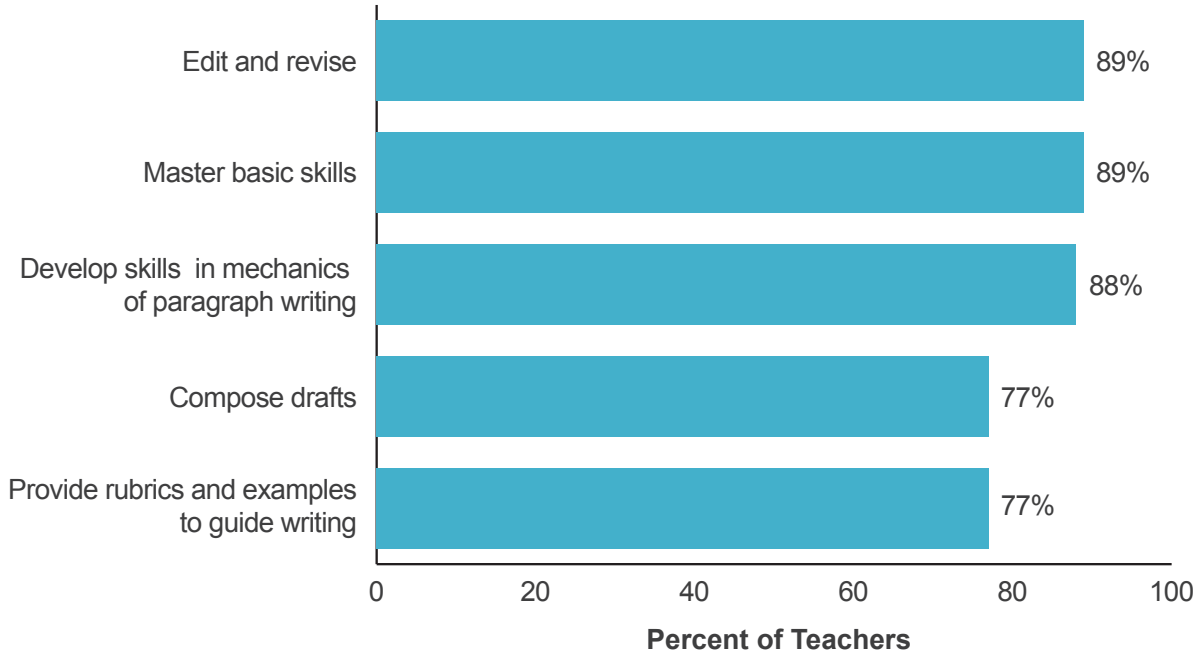
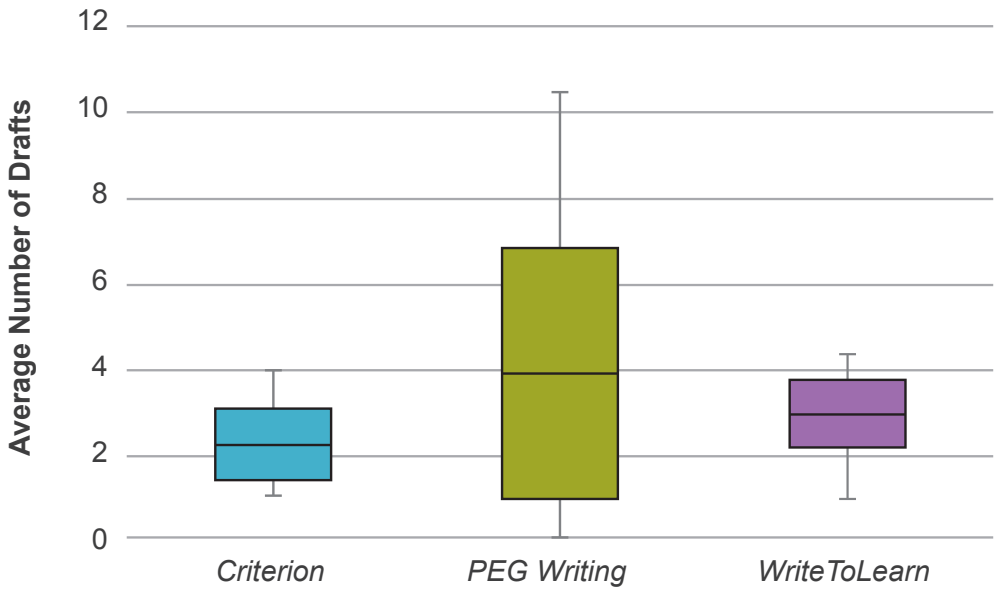


Exhibit 14: Average Number of Drafts Per Assignment by Product

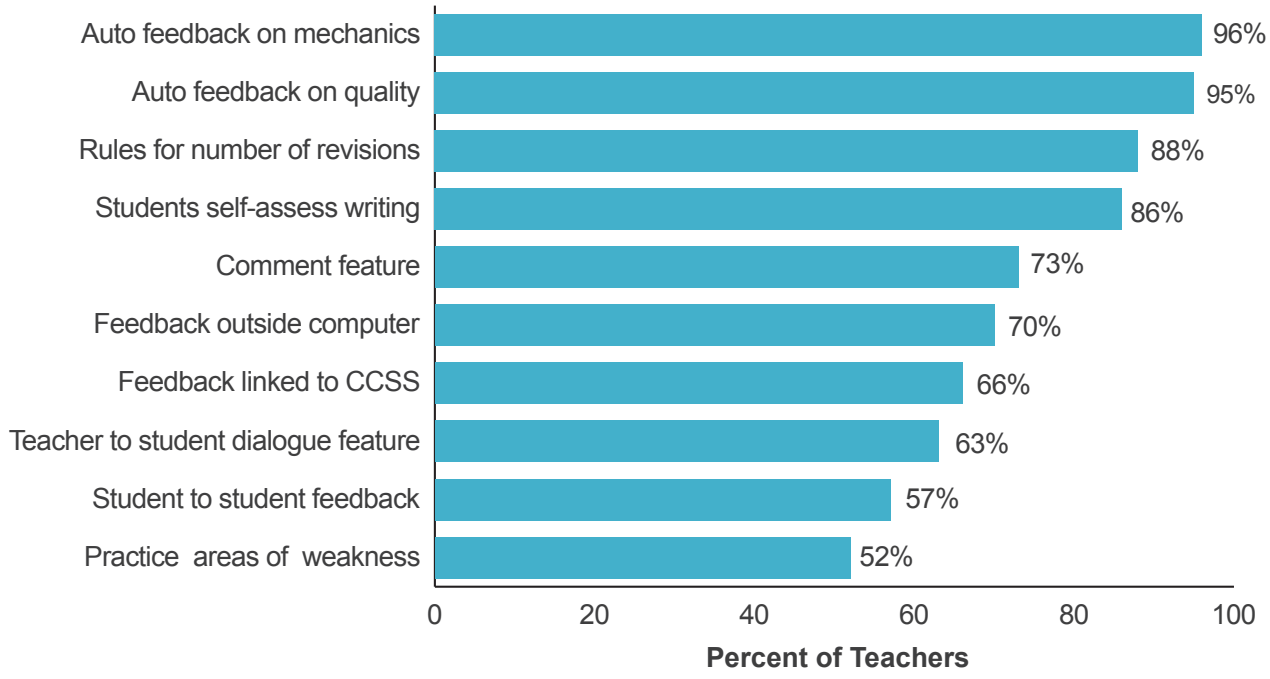


Use of Software to Provide Feedback on Writing

The majority of teacher survey respondents also said the writing software played a moderate to major role in each of 10 subtasks related to providing feedback on student writing, as shown in Exhibit 15. Nearly all software-using teachers employed the software's machine scoring feature to provide students with feedback on grammar, mechanics, organization, and sentence fluency. In addition, more than 60% augmented this automated feedback with their own comments on the content and organization of the writing, using the software's comment (73%) or teacher-student dialog feature (63%). Similarly, half of the LDC teachers who completed logs said they used the software's "comment" feature to provide students with feedback. Of teachers surveyed, 70% also reported that they provided feedback outside of the software; half of those who completed a log also said they provided in-person feedback to supplement the automated scores generated by the software. Hence, **these LDC teachers used the software's automated feedback but did not rely on it as the sole mechanism for giving students information about the quality of their writing.** In interviews, teachers said that they were grateful that the software could provide feedback on grammar, usage, and mechanics, freeing them to focus on more substantive issues regarding the writing content. One teacher using *PEG Writing* described the role of the writing software in providing feedback:

> [*PEG Writing*] *does the part that is time-consuming—the grammar things, the mechanics things, the conventions. I really appreciate that. . .Our last [essay] was about how fairy tales have changed. PEG will know the sentence structure things, but it would never know if you wrote about whether fairy tales have changed or if you wrote about chocolate chip cookies because it doesn't have that capability. I have to read for content to see if you're doing that—if they are citing things correctly, if you have put enough meat and support, and if you're really focused on the prompt.*

Exhibit 15: Teachers' Ratings of the Role of the Software in Providing Feedback on Students' Writing as Moderate to Major



Interviews provided additional information on promising practices some teachers used to ensure that the students were actually learning from the feedback provided by the software product and applying the feedback to their writing. One teacher, for example, who was using *WriteToLearn* with English Language Learners, asked students to print out the feedback provided by the software product, which she then reviewed one-on-one with each student to discuss the edits their summaries needed. In another teacher's class, after students finished a tutorial recommended by *PEG Writing* based on errors in their writing, the teacher required the students to explain the writing concept that they missed and how they planned to revise their essay.

The teachers we interviewed also described using the software for peer review and feedback, although that practice was less widespread than using the software to provide teacher feedback to students—a finding supported by both survey and teacher log data. As shown in Exhibit 15 above, 57% of teachers said that software played a moderate to major role in supporting student-to-student feedback and of the LDC teachers who completed the logs, 33% said that they asked their students to engage in peer review using the software once a week or more over the 4-week reporting period.

Analysis of backend user data on peer reviews from *Criterion* and *PEG Writing*[15] seemed to suggest that **students at these grade levels generally do not provide in-depth feedback on other students' writing.** On average, students using *Criterion* provided far fewer than 1 in-line comment per piece of writing (0.32), and students using *PEG Writing* only completed about 30% of the reviews for their peers that teachers assigned. **Moreover, the feedback provided was short—4.56 words on average.** To better support students in providing more useful feedback to their peers, some of the teachers we interviewed described promising practices they used to improve the quantity and quality of peer feedback (see box).

Promising Practice: Peer Review

A seventh-grade advanced ELA teacher led her students through a "double blind" peer review process. The teacher assigned each student an anonymous peer reviewer in the system. The process then proceeded in the following manner: Student A finished editing his essay and submitted it to peer review. Student B received the essay as a peer review assignment in her interface, provided comments and then resubmitted it to the system. Student A received the feedback in his interface to review. This process took several days, and in her interview, the teacher reported that students were responding well to the practice.

Two other teachers provided additional support to help their students offer higher quality feedback to their peers. One teacher using *PEG Writing* said that she noticed that only her advanced students were providing useful, constructive feedback; others simply said that the paper was good or that the writer did a good job. Consequently, she tried to model feedback and give students guidance on what they should be looking for. She also offered students suggestions for specific phrases that could then be adapted to the paper they were reading. Researchers also observed one experienced teacher in the fall of 2014 providing instruction on how to give constructive feedback. Before students used the peer review feature in *PEG Writing*, the teacher had them write and share feedback on a short paragraph that the teacher wrote. Students then discussed aspects of the feedback that were and were not useful.

[15] *WriteToLearn* (Pearson Education) did not offer a peer review tool at the time teachers used the software.

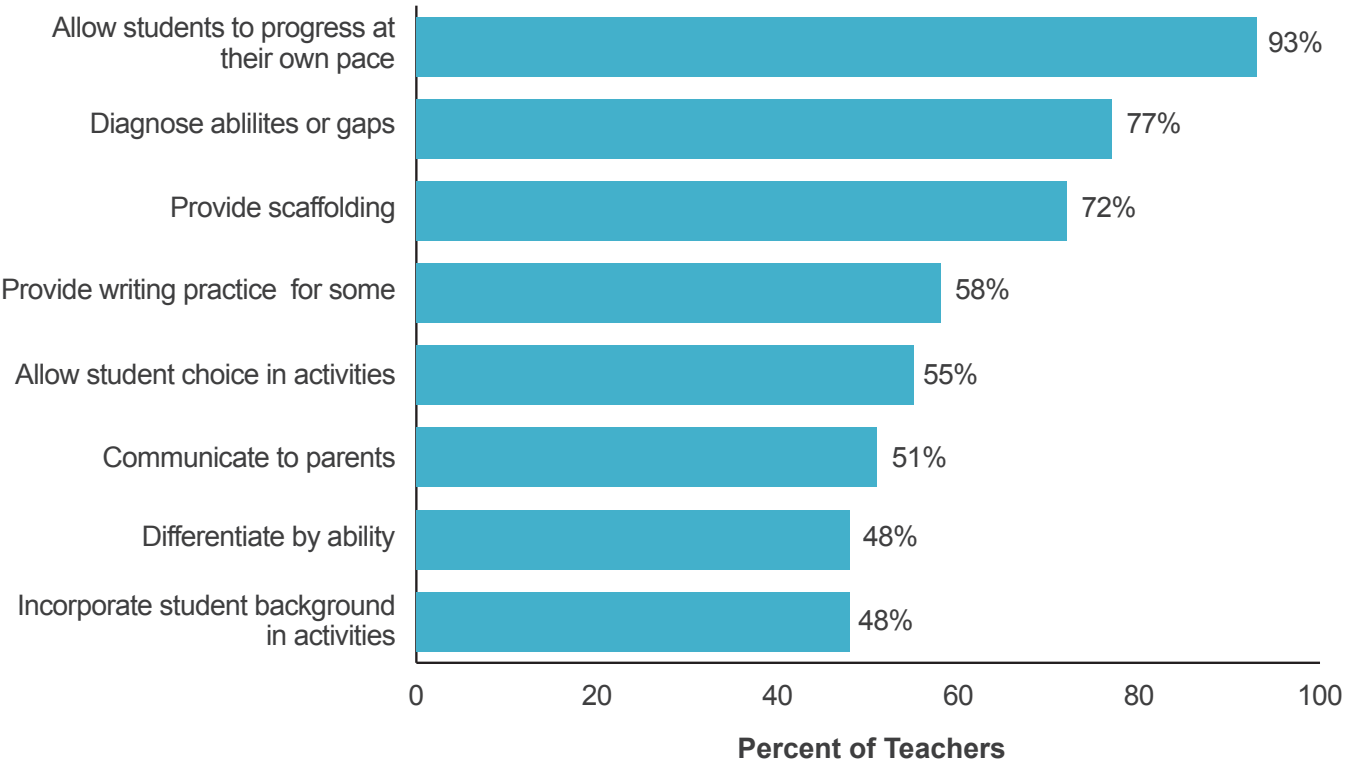
Use of Software to Personalize Learning

A majority of teachers reported that the software played a moderate to major role in scaffolding and diagnosing differences in abilities or learning gaps, and nearly all teachers said the software played a moderate to major role in letting each student progress at his or her own pace on the same assignments, as shown in Exhibit 16.

In interviews, teachers identified ways in which writing software enabled them to personalize learning for their students. One teacher reported that with the data provided by *PEG Writing*, she could regroup her students with other students who were struggling with the same topic much faster than she would have been able to if she had had to wait for a completed essay to diagnose areas for improvement. For example, when she saw students struggling with transitions while they were writing their essay, she asked all students to watch one of the lessons on transitions at the start of the next class before they continued to write. Similarly, another teacher using *PEG Writing* said that the system data helped her group students on the basis of their needs and allowed her to focus her instruction on skills that her students had not yet mastered. Because of competing demands on her time, this teacher reported that without the data provided by *PEG Writing*, she might easily have lost sight of student needs. One teacher using *Criterion* described how the software allowed her to conduct "miniconferences" with students:

> *Because we have so many kids and so little time, it has become impossible to conference with my kids. . .*Criterion *gives me the opportunity where the kids will call me over for something and I can explain something to them and it's almost like a miniconference. . .It helps me differentiate, too,*

Exhibit 16: Teachers' Ratings of the Role of the Software in Personalizing Learning as Moderate to Major



because there are some kids who will focus on, okay I need a comma, but there are some kids who will get deep into the grammar stuff and they will be curious and they will ask, well, what does this mean? It may not be a grammar concept that I have taught before, but I can teach it to them right there, whereas with a kid who is lower, if he were listening to it in class, it would go over their heads.

Two experienced teachers described their strategies for using data from the software to serve students with special needs better and to keep parents updated on their children's progress (see box).

Promising Practice: Using Data from Writing Software

One teacher used the software to inform her special education students' Individualized Education Plans. She found the software useful in allowing her to access data for students in all of her classes. She liked having that data constantly available rather than having to wait for a unit or semester test, and she appreciated the ability to compare scores over time to determine if students were improving.

A teacher at another school thought the system data facilitated his ability to share information about his learning disabled students' writing with their parents. The software creates a portfolio that shows individual student progress and areas that a student continues to struggle with. He shares these student portfolios with parents, who he said appreciate this information and their ability to go into the system themselves to see how their children are progressing.

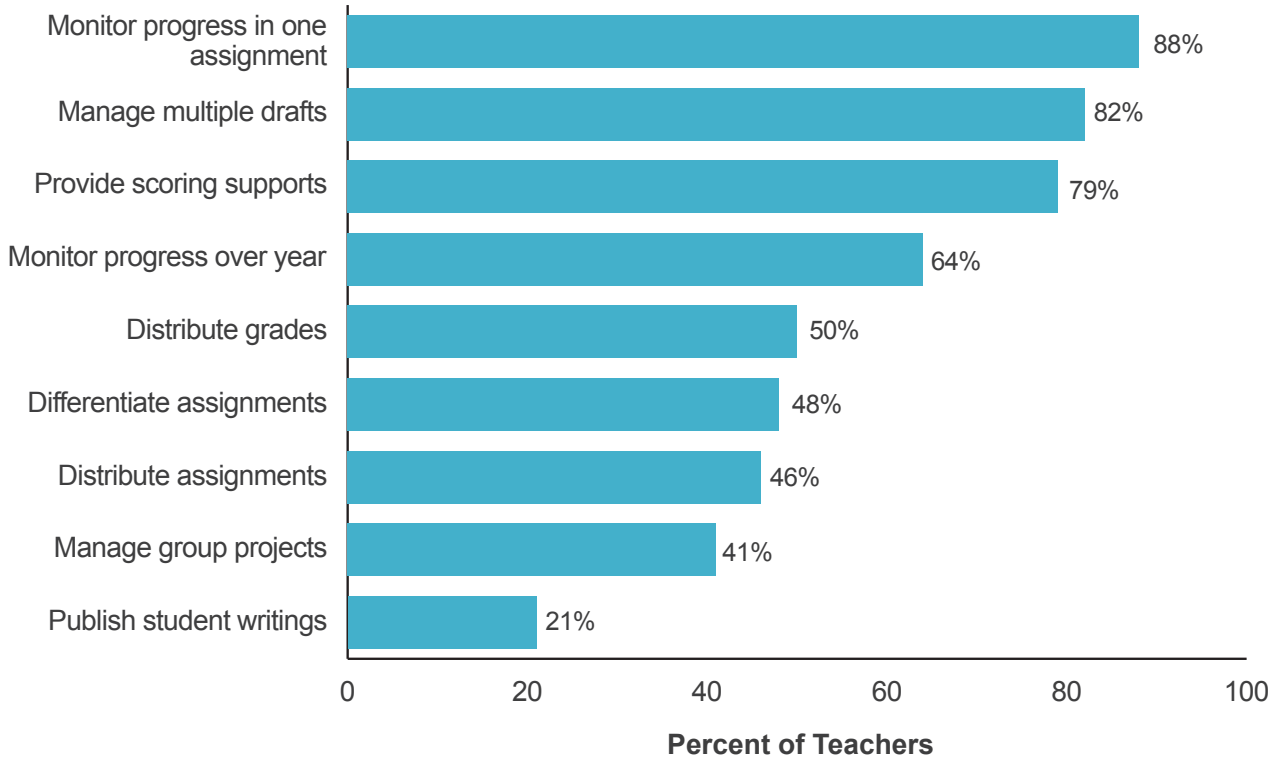
Fewer than half of the teachers surveyed said that the writing software played a moderate to major role in providing opportunities to incorporate knowledge or interests related to students' backgrounds into their writing or providing different writing assignments for students of different ability level.

This finding aligns closely with data from the teaching logs. Teachers who completed the logs reported that they were more likely to use the writing software to provide the same activities for all students (67%) than to allow different students to undertake different activities (33%). Some of the LDC teachers interviewed noted that they could not use the software to easily assign different activities for different students; thus teachers may not have been able to or have known how to differentiate activities in these ways even if they had wanted to do so.

Use of Software to Manage Writing Assignments

Teachers most commonly reported that software played a moderate to major role in monitoring student progress over the course of one assignment (88%), managing multiple drafts (79%), and providing scoring supports (79%). The LDC teachers completing the teacher logs most commonly used the software to track student completion of assignments (75%). As shown in Exhibit 17, very few teachers reported using the software to publish student writing (21%).

Exhibit 17: Teachers' Ratings of the Software Roles in Managing Writing as Having Moderate to Major Effects



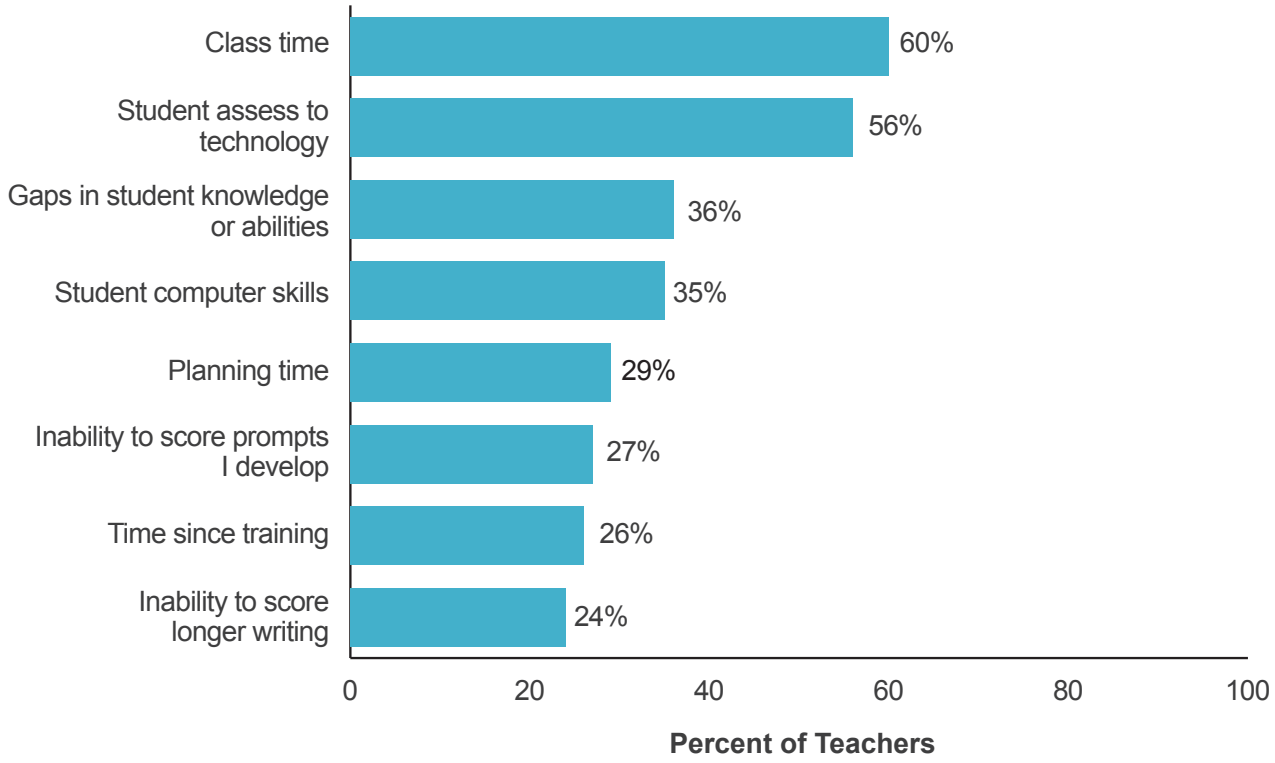
Challenges in Using the Software

Challenges for implementing technology in K-12 classrooms have been studied extensively. The Classroom Trials teacher survey asked about issues that have emerged as challenges in earlier research, including lack of time in the instructional schedule for the technology-based activity, lack of the needed technology infrastructure, lack of needed teacher training, and incompatibility with other mandated instructional approaches.

Survey respondents were asked to rate the extent to which a number of potential challenges had an impact on their use of the writing software (Exhibit 18). **More than half of all LDC survey respondents said that insufficient class time (60%) and limited student access to technology (56%) moderately to significantly reduced their use of the writing software.** In interviews, teachers reported that relying on computers in common labs, libraries, or media centers limited the amount of time they could access the computers and use the software, particularly as the time for standardized testing approached. Teachers had to share computers with other teachers who were beginning to prepare their students to take these tests online. **Approximately a third of the teachers interviewed specifically cited this lack of access to computers as a challenge to using software on a more consistent basis.**

The amount of class time available and competing priorities are the main limitations on implementation teachers have cited in nearly every study of educational innovations. What is perhaps more surprising is that access to enough computers for every student to write remains a significant challenge in these schools. Across all software products, there were few significant differences in the barriers cited.

Exhibit 18: Teachers' Ratings of Various Issues as Having Moderate to Significant Effects on Software Use



Other challenges posed a moderate or significant challenge for smaller proportions of teachers. About a fourth, for example, considered the software's inability to score the prompts they developed (27%) and to score longer writing assignments (24%) as moderate to major challenges. Scoring was of a particular concern for some of the teachers interviewed. These teachers wanted their students to use the software to write essays for LDC modules, but the software offered limited feedback on prompts that were not prepopulated in the system or no feedback at all on the longer essays encouraged by LDC as these essays exceeded the software's word limits. In addition, several teachers said that the scoring was too easy, especially so because the software could not score content; others worried that the scoring was too hard for their less advanced students. Another *PEG Writing* user was unhappy that the software could not judge weak versus strong evidence or identify plagiarism.

One potential challenge noticeably absent in the great majority of teachers' responses was technical difficulties with the software. Overall, 71% of teachers who used the software said they experienced technical problems "never" or "rarely." Among those who did experience technical problems, respondents said they were rarely so serious as to constitute a barrier to using the software. For teachers who did experience technical problems, software glitches or bugs (38%) and students' computer malfunctions (38%) were the most common, followed by unreliable Internet (31%) and insufficient bandwidth (22%). There were no significant differences in the prevalence or impact of technical problems in regard to the software product used.

Teacher reports of barriers to use did not vary significantly in regard to software product, as shown in Exhibit 19. Even though most teachers across all three products were trained at about the same time near the beginning of the school year, *PEG Writing* users were significantly

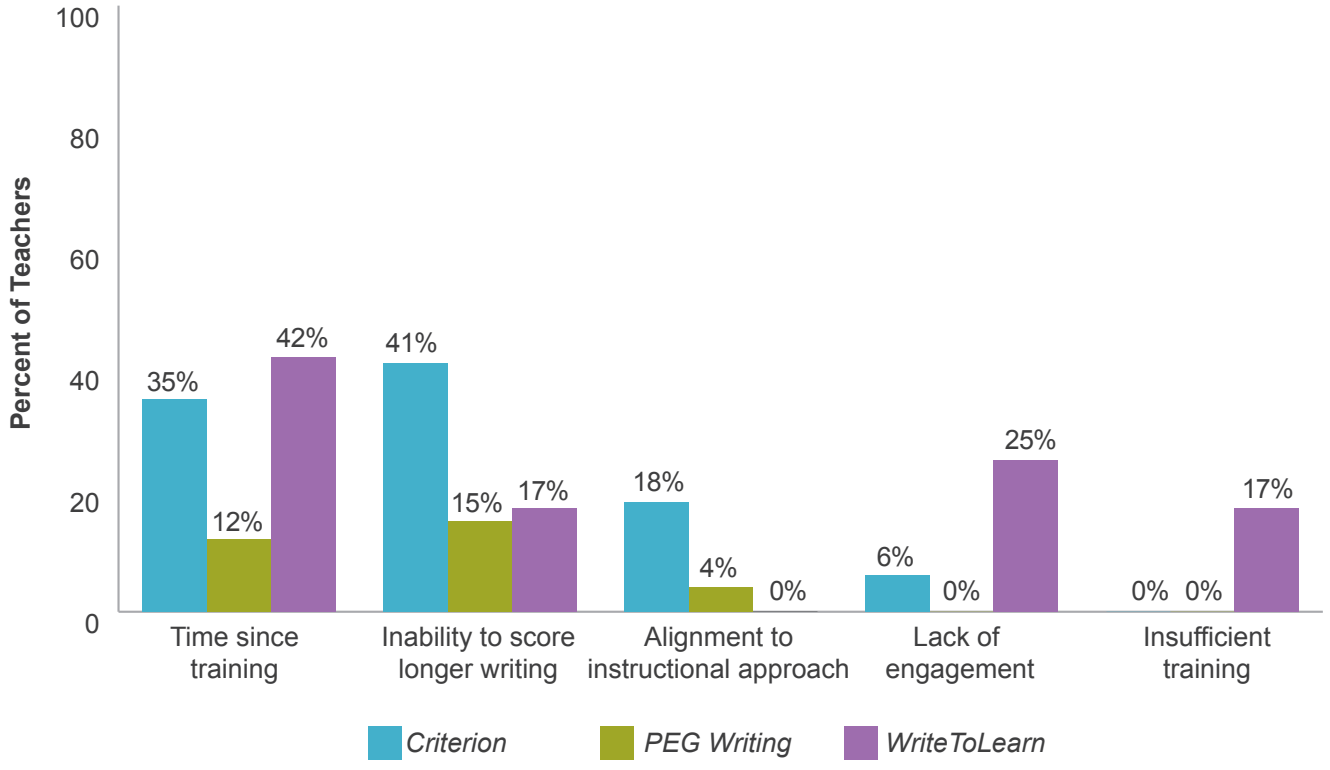
less likely to indicate that the time since their training had a moderate to significant effect on their use of the software ($p<0.05$). On the other hand, *Criterion* users were more likely to say that inability of the software to score longer writing samples and lack of alignment with their own instructional approach were barriers to using the software ($p<0.05$). *WriteToLearn* users were significantly more likely to cite lack of student engagement and insufficient training as having a moderate to significant impact on their use of the software ($p<0.05$).

Challenges for Non-Users

Of the LDC teachers trained to use the software products in the study, 20% of those surveyed reported that they had not used the software. The key challenges to use that they cited were very similar to those identified by the software users. Most often, the nonusers cited the barriers of a lack of time to fit writing software into their curriculum (46%), followed by lack of technology access (38%).

Access to computers may have been a factor in these teachers' decision not to use the software. Although a greater proportion of nonusers than users said they had access to at least some computers in their classroom (62% compared with 25% of software users), users of writing software were more likely to have a 1-to-1 ratio of students to computers (80% compared to 67% for nonusers). However, this difference in the likelihood of use between teachers with and without access to 1-to-1 computing for their students was not statistically significant. There were no significant differences between users of writing software and nonusers in teachers' self-reported technology skill level with computers.

Exhibit 19: Variation in Barriers Rated as Having a Moderate or Significant Impact on the Use of Software, By Product



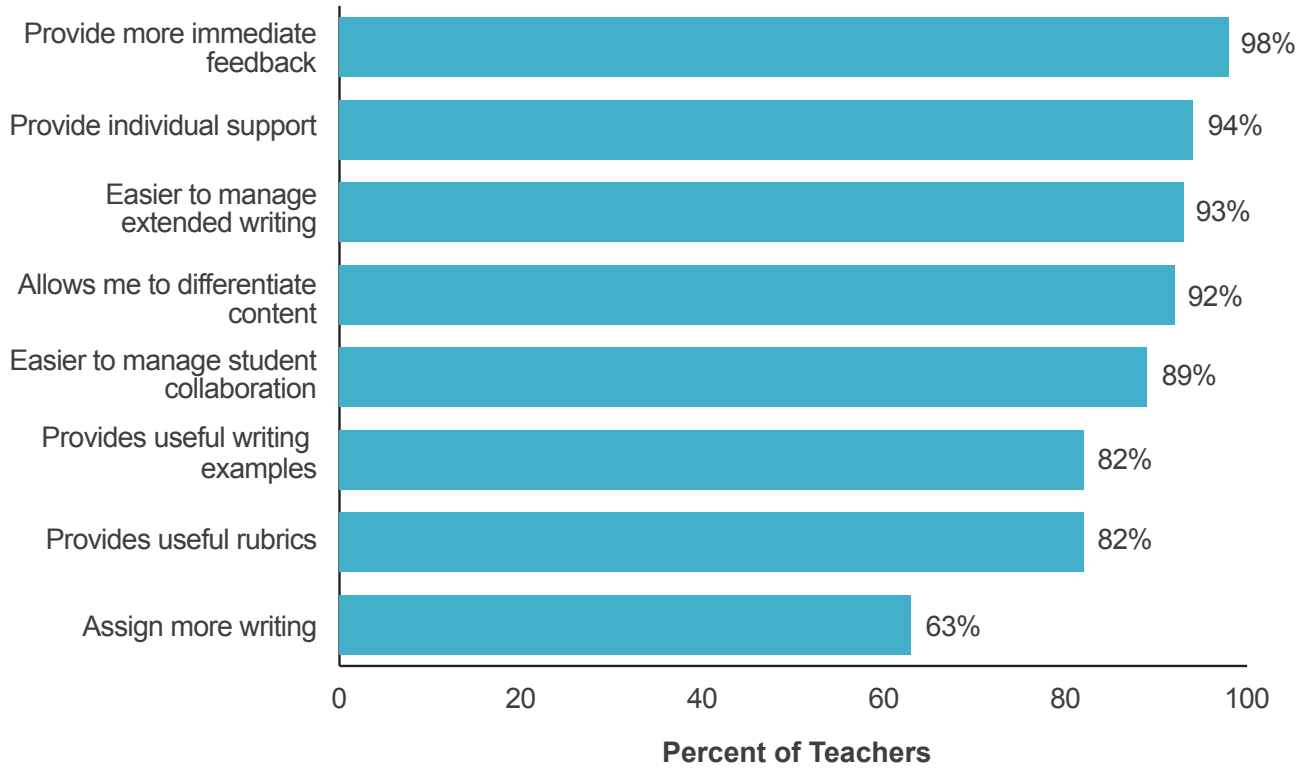
Teachers' Satisfaction with the Writing Software

LDC teachers had favorable impressions of the writing software and the effect of the software on their writing instruction. As shown in Exhibit 20, a large majority of the survey respondents (more than 80%) agreed or strongly agreed that the software improved their ability to:

- Provide more immediate feedback to their students
- Provide individualized supports
- Manage extended writing assignments
- Differentiate content
- Manage student collaboration
- Provide useful writing examples
- Provide useful writing rubrics.

In addition, **more than three out of five teachers agreed or strongly agreed that the software led them to assign more writing,** in particular extended writing pieces. One *WriteToLearn* teacher, for example, said, "Compared to last year, I didn't assign as many writing tasks in the form of paragraph writing." A *Criterion* teacher reported that he went from assigning no prompts at all last year to assigning "a lot more this year."

Exhibit 20: Teachers who Agreed or Strongly Agreed with Various Benefits of Using Software in Writing Instruction



Expanding on the survey reports, nearly every teacher interviewed said that he or she found that the instantaneous feedback on spelling, grammar, and other writing conventions the software provided to students—an aspect of grading that is often time-consuming—to be one of the software programs' most valuable features; that feedback gave teachers time to focus on other aspects of student writing. One *WriteToLearn* teacher, for example, reported that he could now grade 115 essays in just one day because he did not have to focus on spelling, grammar, and other writing conventions.

Finally, teachers reported in interviews that the **writing software helped them meet the needs of all students better, which was often a challenge with larger classes.** One *PEG Writing* teacher noted that writing software's ability to meet students' individual needs:

> *It [the software] is pushing the higher kids even higher, and it's meeting the kids where they are and pushes them from where they are. It's a challenge to get to 100 or so kids to push them. I do that through writing conferences, knowing where they are, and trying to push them through my notes, but PEG does it so much faster.*

A teacher using *WriteToLearn* echoed this sentiment, "It's made teaching more personalized. . . See how that student is looking at her score. . . It makes it easier for them to correct their own mistakes. If they get a wrong the second time, I can step in."

> *"Having that quick feedback right when they submit it to see what they can change, without that program, we wouldn't have spent the time to really go through those things. That would have taken forever to do. It has helped to speed up that editing and revising process."*
>
> – Grade 6 ELA teacher

Teachers' Perceptions of Effects on Students

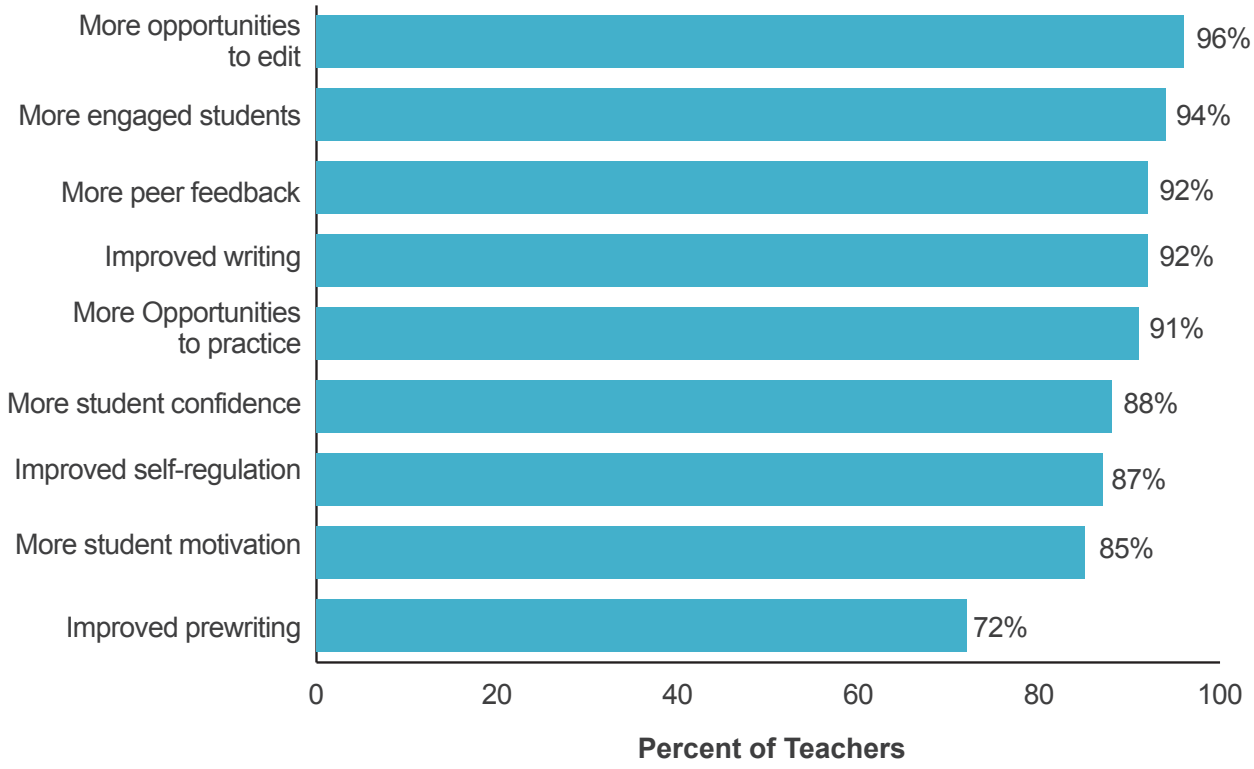
In surveying and interviewing both experienced and LDC teachers, we asked them to reflect on how the writing software products affected student outcomes, including benefits for their students and for their teaching. Although some variation occurred by product, overall the teachers concluded that students were positively affected and that many benefits accrued from using writing software in their classrooms.

Changes in the Students' Writing Activities

Teachers were generally positive about the effect of the writing products on student outcomes. As seen in Exhibit 21, **nearly all teachers surveyed agreed or strongly agreed that the software gave students more opportunities to edit their work, improved student engagement, encouraged peer-to-peer feedback, improved students' writing, and gave students more opportunities to practice writing.**

Both LDC and experienced teachers across all three products cited the appeal of the "gaming" aspect of *Criterion*, *PEG Writing*, and *WriteToLearn*. Students liked receiving writing scores immediately and then

Exhibit 21: Teachers who Agreed or Strongly Agreed that Software Use Had an Impact on Student Outcomes



trying to beat those scores with each new draft they submitted. This aspect of the product appeared to motivate even reluctant writers to revise their essays. In addition, 5 of the 7 *PEG Writing* teachers interviewed noted that students enjoyed being able to use the computer to complete their assignments and produced longer writing pieces than those they would have written using a paper and pencil.

Improvements in the Quality of Student Writing

The quality of student writing also improved according to many of the teachers interviewed. To try to raise their scores, students spent more time revising their essays, which teachers said they had not been as willing to do in the past. A teacher using *Criterion* and one using *WriteToLearn* both noted that the software made students slow down and be more thoughtful about what they were writing.

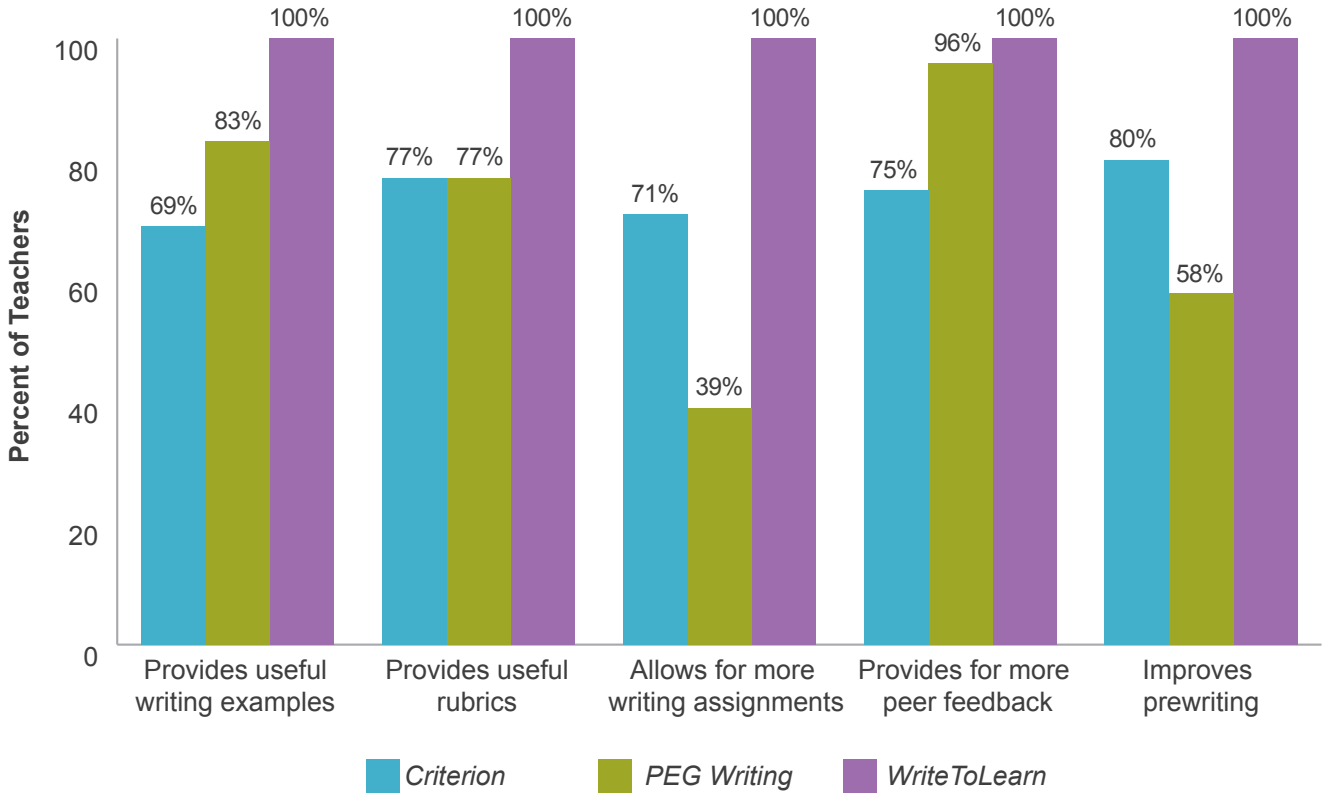
> *"They [students] are more conscious of what the topic is asking them, where before they would have skipped over that and about using evidence as well. They know that in WriteToLearn, it asks them to provide the evidence from the passages and if they are missing that component, it will know that they have gotten that wrong and that they need to go back and fix it."*
>
> – Grade 6 ELA teacher

Teachers who used *WriteToLearn* were universally positive in reporting that the software had led them to assign more writing, provided useful writing examples or rubrics, improved students' prewriting, and increased peer feedback; teachers who used *PEG Writing* and *Criterion* were slightly more reserved in their assessments. As shown in Exhibit 22, *Criterion* and *PEG Writing* teachers were significantly less likely to agree or strongly agree that the software provided useful writing examples or rubrics, led them to assign more writing, or improved students' prewriting ($p<0.05$). *Criterion* teachers were also significantly less likely than *PEG Writing* and *WriteToLearn* teachers to agree that the software provided more opportunities for peer-to-peer feedback ($p<0.05$). This finding may be attributable to *Criterion* teachers reporting that the software played a smaller role in providing feedback in general compared to the other two systems.

The *WriteToLearn* teachers were more likely to work with lower-ability students, which may explain why their perceptions of the software's effect on students were so much more positive than those of teachers using the other two products. However, the survey response rate for *WriteToLearn* teachers (51%) was lower than those for *PEG Writing* and *Criterion*. Thus, the *WriteToLearn* survey respondents may have been the more enthusiastic members of this group and thus may constitute a biased sample when all *WriteToLearn* users are taken into consideration.

Although there were no other significant differences in teachers' reports on the effects of using writing software, either by grade level or the types of students taught, several teachers interviewed noted that students of different ability levels interacted with the software in different ways. One *Criterion* teacher

Exhibit 22: Variations in Teachers who Agreed or Strongly Agreed that Software Use Positively Affected Instructional Practices and Student Writing, by Product



said that because lower-level students could quickly see their mistakes and how they could fix them, they felt they were improving their writing. For higher-level students, who were confident about their writing at the outset, *Criterion* often revealed weaknesses in their writing that they were determined to fix. A teacher using *PEG Writing* noted that the software built the confidence of struggling writers:

> *It builds their self-esteem as far as what they can accomplish when they are using PEG. A lot of them never thought that they could accomplish as much as they do with their writing with PEG. Definitely a boost for a lot of them, especially in my low-level class. All my classes enjoy it, they all enjoy seeing their grade go up. . .but for my kids that start out with [receiving] 1 point [on each of the six traits of writing—ideas, organization, voice, word choice, sentence fluency, and conventions], they go up 1 point in one category, and they are excited.*

Finally, favorable attitudes toward the writing products teachers had used was reflected by **two-thirds of survey respondents indicating that they were "very likely" to recommend the writing product to a colleague.**

Teacher Recommendations for Additional Supports and Program Improvements

All LDC teachers were trained to use the writing software during a half-day training session in fall 2014 in keeping with the vendors' typical training models. The experienced teachers participating in Classroom Trials pilot work also reported receiving an initial training session before using the writing software. However, unlike the experienced teachers, LDC teachers received little regular, ongoing training and support, either from the vendors or through their participation in LDC. Experienced teachers tended to come from schools or districts that had opted to purchase and implement the writing software products, and a staff member, usually an academic coach, served as support provider. The LDC teachers did not have this kind of support. A few LDC teachers reported emailing or calling the vendor to ask technical questions, such as how to set up class lists.

Across all three products, LDC teachers interviewed said that they would like to have had more ongoing support. One teacher using *PEG Writing* said that she would have liked another training session to learn more about features that were only briefly covered during the initial training and are used less often, such as the peer review tool or the minilessons. Another teacher noted that she would have liked to have had a *PEG Writing* representative in her classroom early on to provide immediate support.

In general, the teachers we interviewed were comfortable with learning how to use the writing software, either through practice or by receiving help from colleagues. Several teachers noted modifications to the training that would have been useful at the outset, including more information on

Teacher-recommended Software Refinements

- One *PEG Writing* user recommended moving the feedback on one draft to subsequent drafts so that teachers can more easily see whether students have addressed the feedback.
- Another user would like *PEG Writing* to allow students to keep working on their draft while peers provide feedback.
- A *WriteToLearn* teacher recommended that the software show student numbers when teachers are setting up the class roster at the beginning of the year and allow her to track students who have and have not completed the assignment.
- Teachers using *WriteToLearn*—the only product with a summary feature—wanted a greater variety of texts in languages other than English and Spanish, as well as those for students at lower reading lexile levels.
- *Criterion* teachers said that they would like their students to be able view their prewriting on a split screen while they are writing so that they could easily refer to both documents at the same time.
- Two of the *Criterion* users recommended allowing teachers to adapt the software's pre-writing tools based on the writing prompt. One user, for example, noted that she could not a find a tool that allowed students to write down their counterarguments while planning for an argumentative essay.

how to use the writing software in settings with limited computer access, how to use specific features (e.g., peer review), how to use the formative data, and more time to actually use the product during training.

In addition to training, when asked about additional supports, teachers most commonly provided suggestions for how the writing software products could be modified or improved to make them more useful in the classroom. **Across all products, teachers wanted more flexibility to adapt the products and their features to the unique needs of their students.** For example, one of the teachers using *Criterion* said that she would like for students to be able to create assignments to practice writing on their own or for other classes. Two teachers using the *WriteToLearn* software noted that they would like to be able to give students different reading and writing assignments depending on student ability. *PEG Writing* teachers said that they would like to be able to tailor the feedback automatically provided to students.

Several teachers reported challenges in using the writing software products with lower achieving students and recommended software changes to meet the needs of those students better. For example, one teacher suggested allowing teachers to tailor their scoring and feedback for lower achieving students. Another recommended adding an option for unscored drafts so that struggling students would not be discouraged when they received a "0." In addition, teachers using all three products recommended changes to the software to make it more convenient and user-friendly for both teachers and students.

Compatibility of the Writing Software with the CCSS and LDC Approach to Literacy

For teachers implementing curriculum modules developed through LDC, the writing software was typically used at the end of the process when students had to write essays. For example, one teacher of English Language Learners said that when she taught an LDC module, she typically had students complete a reading, followed by a group discussion using the Socratic seminar approach; then students wrote their response to a prompt using the software as the concluding activity.

Teacher interviews suggested several areas where the compatibility of the writing software with the LDC approach to literacy instruction could be improved. **LDC emphasizes evidence-based writing and connecting writing to research, and a number of teachers felt the software products need to do more to support students in that kind of work.** Teachers wanted more support for the conventions of evidence-based writing, including more minilessons on how to properly cite research and create bibliographies. Teachers also noted that the products often could not verify whether citations were done properly or treated them as grammar errors.

A *Criterion* and *PEG Writing* teacher both noted that they would like to see greater alignment between the LDC rubric and the one used by *Criterion* and *PEG Writing*. One *WriteToLearn* teacher also said that they wanted to easily insert their own prompts and reading materials and have the system provide the full range of rubric scores for writing to these prompts (a capability not currently offered by the existing system, which must now be "trained" using a set of student writing samples). Finally, several teachers also wanted students to be able to view their prewriting products and their unfolding composition at the same time so that they could move back and forth between the two.

Conclusion

This exploratory study provides a first look at how three writing software products can be used to support teaching and learning of the challenging CCSS in ELA/Literacy. The major assumption underlying these Classroom Trials was that middle school teachers trying to implement CCSS would view writing software products as a useful addition to their toolkit. Our findings with respect to this assumption were clearly affirmative: **94% of teachers said their students were more engaged in writing when using the software, and 92% said their students' writing improved as a result of using the software.** Two-thirds of LDC teachers said they were "very likely" to recommend the software product they had used to a colleague. This finding stands in contrast to some of the negative press that machine-scoring of student writing has received among writing teachers (e.g., see the National Council of Teachers of English's Position Statement on Machine Scoring, 2013).

We conclude that the positive responses of teachers in the Classroom Trials to the writing software, including its automated feedback feature based on machine scoring of student writing, is the nature of their implementation model. Classroom Trials teachers were volunteers, and there was never any doubt that they were in control of the writing instruction in their classrooms: They were free to decide how and how frequently they would have their students use the writing software. **Classroom Trials teachers regarded the software as a potential amplifier of their writing instruction efforts rather than as a replacement for their judgment and skills as teachers.**

The ways in which teachers used the writing software were first and foremost to support students as they were writing and to provide feedback. These two instructional tasks were very intertwined because the immediate feedback the system could provide on overall writing quality and writing mechanics were strong influences on students' revision processes.

At the same time, the LDC teachers were not naïve about the limitations of automated scoring of student writing. They understood that the automated feedback would not necessarily be a good indicator of how well the student really understood the material he or she was writing about. But LDC teachers were willing to accept that limitation because the time that the writing software saved them by providing feedback on writing mechanics could be spent focusing on the content of students' writing. These LDC teachers were developing instructional practices built on a human-machine partnership, seeking to capitalize on their relative strengths in different areas.

A major goal of the Classroom Trials was to identify promising practices for use of writing software in support of teaching the CCSS ELA/Literacy standards. Most of the LDC teachers in the Classroom Trials appeared to be using a trial-and-error approach to developing instructional practices incorporating the writing software. In some cases they had access to colleagues who were also using the product, but that was not always the case. Teachers looked for matches between principles of writing instruction they had already assimilated (e.g., individualizing instruction for a particular student) and capabilities of the writing software. Given that these LDC teachers were trying out the software for the first time, further innovation is expected as greater experience is gained. Ideally, communities would emerge of teachers using the same writing product and teaching similar students; they would share tips, writing prompts, and strategies.

Practices that SRI researchers judged to be worthy of emulation and further research included:

- Using software-generated rubric scores to identify common writing weaknesses that can be addressed in teacher-led lessons before students continue work on their essays
- Using system data to group students with similar skill needs for small-group work on their common weakness
- Having students complete a reading, engage in a group discussion of the reading content, and then write in response to a prompt related to the reading using the software
- Modeling informative feedback on writing so that students can determine what to look for and how to provide feedback to their peers that actually informs revisions
- Developing a collection of starter sentences that can scaffold students' efforts to frame constructive feedback on their classmates' writing
- Double-blind peer feedback, with students exchanging their writing and scoring others' work without knowing their identities
- Using the software to create a portfolio of each student's writing and associated rubric scores over time to illustrate individual student progress and areas for further work.

Teachers developed these promising practices through their own use of the writing software. Neither the vendors nor the schools and districts set up ongoing support for the writing software users after their initial half day of training on the system. It seems likely that intentionally creating a community of users using writing software with similar students would have greatly accelerated the emergence and spread of best practices. Additional years of experience with an instructional software system can also be expected to lead to more effective practice (Pane, Griffin, McCaffrey, & Karam, 2014).

The ambitious, cross-curricular extended units that LDC teachers sought to implement certainly put more stress on the writing software platforms than more conventional writing instruction would have. In this context, it is not surprising that LDC teachers wanted to change several aspects of the systems. Better capabilities for supporting the input of teacher-developed writing prompts and associated research materials without losing the automated feedback capabilities would be first on the list of improvements. A number of LDC teachers expressed concern also that the systems could not handle the length of writing assignment to which they aspired. Finally, it also seems clear that the systems' prewriting tools are not yet as easy to use and compelling as they need to be to garner more widespread use. Teachers seemed to feel it was easier and more natural for students to plan their writing using paper-and-pencil tools, and they pointed out the ineffectiveness of a system that did not allow a student to view her prewriting material and her emerging composition at the same time.

In closing, we acknowledge the limits of this exploratory study. The teacher samples were not representative of writing teachers nationally, and they were not equivalent across the three products. Nor did this study include the collection of student writing assessment data. The data reported here should not be interpreted as demonstrating the efficacy or lack thereof for any of the three products. Instead, we urge the reader to use our data as a platform for generating hypotheses for how writing software, teacher training and support, and classroom practice might be improved to expand the capacity needed to help students acquire fundamental writing and literacy skills.

References

Attali, Y. (2004). *Exploring the feedback and revision features of Criterion.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school.* Washington, DC: National Academy Press.

Bryk, A., Sebring, P, Allensworth, E., Luppescu, S., & Easton, J. (2010) *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: The University of Chicago.

Gallagher, H., Penual, W., Murphy, M., Bosetti, K., Shields, P., Toyama, Y., & Yee, K. (2009). *National evaluation of writing project professional development, year 2 report.* Menlo Park, CA: SRI International.

Gallagher, H., Woodworth, K., Bosetti, K., Cassidy, L., McCaffrey, T., Yee, K., Wang, H., Shields, P., Murphy, R., & Penuel, W. (2011). *National evaluation of writing project professional development, year 4 report.* Menlo Park, CA: SRI International.

Graham, S. (2008). *Effective writing instruction for all students.* Wisconsin Rapids, WI: Renaissance Learning.

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing, 26*(1), 1-15.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high school—A report to Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers learning: Enhancing opportunities, creating supportive contexts.* Committee on Strengthening Science Education through a Teacher Learning Continuum. Board on Science Education and Teacher Advisory Council. Washington, DC: National Academies Press.

Graham, S., & Sandmel, K. (2011) The process writing approach: A meta-analysis. *The Journal of Educational Research, 104*(6), 396-407.

National Council of Teachers of English (2013). *NCTE position statement on machine scoring: Machine scoring fails the test.* Retrieved from http://www.ncte.org/positions/statements/machine_scoring.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Education Evaluation and Policy Analysis, 36*(2), 127-144.

Preston, D., & Goodman D. (2012). *Automated essay scoring and their repair of electronics.* Retrieved from http://snap.stanford.edu/class/cs341-2012/reports/03-Preston_cs341_-_Dan_and_Danny_-_Final.pdf

Shermis, M. D., Burstein, J. C., & Bliss, L. (2004). *The impact of automated essay scoring on high stakes writing assessments.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Shermis, M. D., & Hamner, B. (2013). Handbook of automated essay evaluation. In M. D. Shermis & J. Burstein (Eds.), *Contrasting state-of-the-art automated scoring of essays* (pp. 313–346). New York, NY: Routledge.

Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies: An International Journal, 3,* 22-36.

Wilson, J., Olinghouse, N. G., & Andrada, G. N. (2014). Does automated feedback improve writing quality? *Learning Disabilities: A Contemporary Journal, 12*(1), 93-118.

Zumbrunn, S., & Krause, K. (2012). Conversations with leaders: Principles of effective writing instruction. *The Reading Teacher, 65*(5), 346-355.

Appendix A. Methods

Sample Selection

Researchers created two contrasting sample of writing teachers: (1) teachers who had been using the writing software before the Classroom Trials study took place, referred to as "experienced teachers"; and (2) teachers working with LDC who had not used the writing software before our study, referred to as "LDC teachers."

Experienced Teachers

Nominated by the software vendors, the sample of experienced teachers was drawn from those who had been using the three products for some time. The vendors based their nominations of 4-8 teachers on reports from their sales representatives, who had the most contact with users, as well as information reported by schools and districts.

We selected one nominated teacher for each product to interview and observe in winter 2014 to pilot our data collection instruments. One vendor, ETS, was unable to recruit teachers in enough time to participate in the pilot visit. The remaining nominated teachers were asked to participate in the spring 2014 site visit and to complete a record of their writing instruction. We also analyzed software backend user data from their students' use of the product. The total number of experienced users who participated in our study is shown by product in Exhibit A-1.

The three *Criterion* (ETS) users came from one private school in Tennessee. For teachers working with Measurement Incorporated, the sample of experienced teachers came from four districts in North Carolina. Although *PEG Writing* was the selected writing software as part of the TCP competition, these districts were using NC Write, which has the same functionalities and scoring engine as *PEG Writing*. *PEG Writing* was in its first official year of use in 2013-14, and thus did not yet have experienced users. *NC Write* also was not part of a contractual agreement with a state or private agency like some of the other Measurement Incorporated writing software products. The teachers who used *WriteToLearn* (Pearson Education) came from two districts in Georgia.

LDC Teachers

To facilitate recruitment for the study and understand how LDC teachers might perceive the value of the writing software products, in March 2014 the study team held an informational session on the three products in a district in Pennsylvania that had a long-standing history with LDC. Volunteer middle and high school teachers across the content areas participated in this half-day session (for which they received a stipend). During this session, teachers were assigned one of the three products to learn about, and each of the vendors made a brief introduction and held

Exhibit A-1. Experienced User Sample by Writing Software Product

	Criterion **(ETS)**	***PEG Writing*** **(Measurement Incorporated)**	***WriteToLearn*** **(Pearson Education)**
Pilot Teacher	0	1	1
Spring 2014 Participant	3	5	5
Total	**3**	**6**	**6**

a training session for its group. Then LDC teachers were given the opportunity to share their thoughts on the products and how they might facilitate the LDC way of teaching. LDC used the feedback in its recruitment efforts of districts for the study, and vendors learned how they could better support teachers in using the writing software.

LDC then recruited writing teachers in districts it was already working with to participate in the main Classroom Trials data collection effort. Recruiting was done in districts in three geographic regions, with one of the writing software products offered for implementation in each region. Teachers in these districts were invited to participate in a training session on the assigned writing software product held in their region in fall 2014. The study design called for training 40 middle school teachers per product at these sessions. Because recruiting proved difficult in the Pennsylvania region, a small number of high school teachers was accepted to enlarge that region's sample; even so, the final available sample (26 teachers) for Criterion fell short of the target.

We then asked all teachers trained in one of the products to complete a survey in December 2014 and January 2015, and using survey data on the frequency of the use of the product since the training, we selected 4-8 teachers per product to participate in a site visit and complete an online, daily instructional log. Our final site visit and log sample included 7 teachers using *Criterion* (6 middle and 1 high school), 7 using *PEG Writing* (all middle school), and 4 using *WriteToLearn* (all middle school), for a total of 18 teachers.

Data Collection Activities

Data sources for this project comprised background research on the three products, visits to both experienced teachers' and LDC teachers' classrooms to conduct interviews and observations, instructional logs collected from experienced and LDC teachers, surveys of LDC teachers after their first semester of product use, and backend user data obtained from each of the three products.

Site Visits

Members of the SRI research team conducted site visits in winter/spring 2014 with 15 experienced teachers and in spring 2015 with 18 LDC teachers. During the visits, we interviewed each teacher and observed one of his or her classrooms that was using the writing software. This approach served to connect observed practices with teachers' descriptions of their intents, beliefs, and perceptions, and to understand the contextual factors that influenced software use. For 4 LDC teachers, our data collection was restricted to teacher interviews because scheduling conflicts precluded classroom observations.

During visits with experienced teachers, we asked how they used the products to facilitate their writing instruction and for what purposes. We asked them to describe what they perceived to be best practices in using the writing software, how they learned those best practices, and what helped and challenged them in using these products.

We asked LDC teachers the same questions about how and why they used the writing software products, the challenges they encountered in learning and using them, and the additional supports they thought would be helpful. We also asked how they used the writing software to implement LDC modules.

Teacher Log

All teachers visited by an SRI researcher were asked to complete an online daily teaching log about their writing instruction. The logs asked teachers to report on a relatively small and recent group of lessons for a particular class, enabling them to provide a detailed and accurate description of use. The log entailed a checklist that captured contextual information such as whether teachers used the software that week, their purposes for using the writing software, and how much time students spent working with it. Our

use of log data in conjunction with the backend user data (described below) from the same period helped in understanding teachers' purposes for using the software and their rationales for the observed frequency and pattern of use.

During our site visits to LDC teachers, we explained the logs to the teachers and asked them to complete a log for each of 4 consecutive weeks following the visit. For our experienced teacher sample, we gained access to the schools too late in the school year to enable them to complete logs for 4 consecutive weeks before the end of the school year. Instead, we asked them to complete the weekly log for however many weeks they had left of school and had them complete a one-time log characterizing "typical use" to provide a more rounded picture of their writing instruction. Of 15 experienced teachers, 11 completed both the typical log and at least 1 week of the daily log. Of 18 LDC teachers, 12 completed the daily log for 4 weeks—4 from teachers using *Criterion*, 7 from those using *PEG Writing*, and 1 from a teacher using *WriteToLearn*.

Teacher Survey

We administered a survey to LDC teachers to assess how they used the writing software to support their writing instruction, the features of the software they and their students used, supports they received for using the tools, and their expectations, if any, for using the writing software in the future. Although we originally intended to train and survey 120 teachers, we were able to recruit only 106 teachers for the study. As shown in Exhibit A-2, we invited 100 teachers to complete the survey (after 6 individuals who took the training were removed from the sample either because they dropped out of the project or because we learned they were nonteaching staff). We had an overall response rate of 70%, with higher response rates for the *Criterion* and *PEG Writing* samples. Although the sample size was adequate to capture large differences in regard to how the three products were used, it was not large enough to capture moderate or small differences among products.

Exhibit A-2. Survey Sample

	Overall	*Criterion*	*PEG Writing*	*WriteToLearn*
Sent survey	100	25	40	35
Opened and completed at least one question on the survey	70 (70%)	20 (80%)	32 (80%)	18 (51%)
Used the software (percent of respondents)	56 (80%)	17 (85%)	27 (84%)	12 (67%)

Backend User Data Analysis

We collected and examined backend system data for the three products to determine student use of the tools, such as how frequently they visited the tool, what features of the tool they accessed, and the amount of time they spent using each tool feature. Analyzing product use at the student level allowed us to explore variation within and between classrooms for both experienced and LDC teachers.

We engaged in two waves of backend data analysis. The first involved exploratory data downloads and analyses in spring 2014 using data from students of the experienced teachers. These analyses enabled us to become familiar with the data elements available from each writing product and to refine our processes for the second wave of analysis. Exhibit A-3 shows the variables available by product.

The second wave of analysis using data from the larger LDC teacher sample (i.e., all the trained teachers who used the product) occurred in spring 2015. Using these data in conjunction with teacher survey responses, we explored patterns in students' use of each system in relation to the instructional goals and practices of their teachers.

Exhibit A-3. Backend System Variables by Product

	Criterion	*PEG Writing*	*WriteToLearn*
Number of assignments	✓	✓	✓
Average number of drafts per assignment	✓	✓	✓
Total number of unique days spent per student	✓	✓	✓
Average word count per assignment	✓	✓	✓
Average change in score from first-to-last draft per assignment	✓	✓	✓
Total time spent on assignments per student	✓	✓	✓
Number of teacher dialogs (teacher to student)	✓		
Number of student dialogs (student to teacher)	✓		
Number of student dialogs (student to student)	✓		
Number of teacher in-line comments	✓		
Number of student in-line comments (e.g., peer review)	✓		
Number of teacher messages (teacher to student)		✓	
Number of student messages (student to teacher)		✓	
Average number of tutorials viewed per assignment		✓	
Average time spent on tutorials per assignment		✓	
Average number of peer reviews requested per assignment		✓	
Average number of peer reviews completed per assignment		✓	
Average length of peer review feedback per assignment		✓	
Number of essay assignments			✓
Number of summary assignments			✓

SRI Education™

SRI Education, a division of SRI International, is tackling the most complex issues in education to identify trends, understand outcomes, and guide policy and practice. We work with federal and state agencies, school districts, foundations, nonprofit organizations, and businesses to provide research-based solutions to challenges posed by rapid social, technological and economic change. SRI International is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society.

Silicon Valley
(SRI International headquarters)
333 Ravenswood Avenue
Menlo Park, CA 94025
+1.650.859.2000
education@sri.com

Washington, D.C.
1100 Wilson Boulevard, Suite 2800
Arlington, VA 22209
+1.703.524.2053

www.sri.com/education

Stay Connected