

# The Validity Case for Assessing Direct Writing by Computer

A Pearson Assessments & Information White Paper

Walter D. Way, Ph.D.  
Laurie L. Davis, Ph.D.  
Ellen Strain-Seymour, Ph.D.

April 2008

# PEARSON



## **The Validity Case for Assessing Direct Writing by Computer April 2008**

### **Introduction**

Technology continues to provide opportunities for changing how teachers give instruction and how students learn. For example, a recent survey by the Sloan Consortium indicated that 63.1 percent of responding K-12 school districts currently have students taking either online courses or courses where online delivery of course material is blended with face-to-face instruction. Furthermore, another 20.7 percent of districts indicated that they planned to introduce such courses over the next three years (Piccaino & Seaman, 2007).

These data regarding online courses reinforce the obvious fact that students are increasingly using the computer in their everyday school work. One such application is the use of word processing software for writing. Clearly, as technology becomes a primary vehicle for students to receive instruction and demonstrate learning, the book reports, compositions and other types of expository writing students do will be almost exclusively authored and edited using word processing programs.

In this age of accountability-based education in the U.S., most states include measures of direct student writing as part of their K-12 assessment system. In nearly all state programs, writing is measured by requiring students to provide handwritten responses. As students become facile with word processing tools and the associated strategies for producing and refining their writing, the case for switching from traditional paper-and-pencil writing assessments to computer-administered writing assessments becomes compelling.

In this paper, we present the validity case for online writing assessments. The core of this argument is that as students increasingly use the computer to write in their everyday work, denying them the opportunity to use the computer to take a high-stakes state writing assessment will likely come to have negative validity implications. In addition, instructing students to write by paper is inconsistent with the preparation that students will need for the word processing demands they will likely encounter in future work or continued education efforts.

Our paper is divided into three major parts. In the first part, we review some of the literature regarding computer-based writing instruction and the associated implications for assessment. In the second part of the paper, we discuss some of the issues and challenges associated with assessing writing by computer. Finally, we discuss some questions that might be considered in establishing a statewide computer-based writing assessment.

### **Using Computers to Write: Implications from the Literature**

The benefits of computers in writing instruction are recognized by the public and supported by the availability of computers for children to access both in the home and in school. For example, Belden, Russonello, and Stewart (2007) conducted a survey examining American public opinion on the importance of writing. They concluded that their survey results “reveal a public that believes computers and other new technologies generally have a positive impact on teaching children to write well” (p. 24). They found that the vast majority of parents of school age children indicated that their children had access to computers at home (87%) and school (82%). These results were consistent across ethnic and demographic subgroups and suggested that computers are available during the day for most American children, either at school or home, or both. In this same survey, 61 percent of the overall respondents and even higher percentages of African American and Hispanic respondents (68% and 69%, respectively) indicated that new technologies including computers are helpful in teaching students to write well.

In another survey conducted as part of the 1998 and 2002 NAEP writing assessments, 97% of grade 8 teachers indicated that their students used the computer to compose (Solomon, Lutkus, Kaplan, & Skolnik, 2004). The teachers participating in the survey also reported that word processing positively impacted students’ motivation, the amount of time they spent revising draft compositions, and their overall writing achievement.

In a meta-analysis study of successful writing instruction practices, word processing was identified as one of 11 most effective strategies for improving the writing of adolescents in middle and high schools (Graham & Perin, 2007). Compared with writing by hand, studies have suggested that word processing has a consistently positive impact on the writing quality of adolescents in middle and high schools. Word processing seems to be especially effective in enhancing the quality of text produced by low-achieving writers (Graham & Perin, 2007).

Given the increased emphasis on using computers and word processing software as part of writing instruction and the positive impact that these tools seem to have on the students’ writing, it follows that students utilizing computers on a regular basis are likely to perform better on online writing assessments than they are on paper-and-pencil based writing assessments. Researchers investigating the comparability of scores for paper and computer versions of a writing test administered to eighth grade students in the context of NAEP Writing Online (WOL) study came to this very conclusion:

Thus, it is conceivable that, for a given level of paper writing performance, students with greater computer facility score higher on WOL because they write better on computer than on paper (relative to their less technologically-experienced peers). And, the reason they write better on computer than they do on paper may be because the computer offers them a tool that makes it possible to do so (Horkay, Bennett, Allen, Kaplan, & Yan, 2006, p. 36).

The results of the Writing Online study have impacted the future plans for assessing writing within the NAEP program. Specifically, in 2011 and beyond, the NAEP Writing assessment at grades 8 and 12 will assess computer-based writing (ACT, 2007). This decision appears to be well supported by the current trends in writing instruction at the middle and high school levels.

For policy makers involved with state assessment programs, this evidence implies the need for attention in that traditional paper-and-pencil direct writing assessments may soon become obsolete. The question at this point is not whether or not writing assessments should be offered by computer, but rather how and by when can states transition their writing assessments so that students will have the opportunity to demonstrate their writing ability using the computer?

### **Issues with Assessing Writing by Computer**

For statewide assessment programs, assessing writing by computer introduces a number of issues. We have chosen three in particular that we believe are appropriate for illustration, and discuss them below. One issue concerns the comparability of traditional paper-and-pencil writing assessments and writing assessments delivered by computer. A related second issue concerns differences in the way human raters score handwritten and typewritten essays. A final issue has to do with the interface that is used to deliver online writing assessments. Other issues or concerns related to computer-based writing assessment no doubt exist but are not explicitly discussed in this paper.

#### **Comparability of handwritten and computer-based writing assessments**

The literature comparing computer-written and handwritten essay performance is mixed. Some studies have indicated higher performance on essays entered using the computer compared with handwritten essays (Russell & Plati, 2001; Russell & Haney, 1997). Other studies have suggested the opposite finding, that is, lower performance for computer-based essays compared with handwritten essays (Way & Fitzpatrick, 2006; Bridgeman & Cooper, 1998). These mixed findings may have to do to the keyboarding skills of the students involved in the studies. For example, the researchers conducting the NAEP Writing Online study reported no significant differences in overall mean scores between grade 8 students using word processors and those handwriting their essays. However, the researchers did find that hands-on measures of keyboarding skill were significantly related to writing assessment performance, so that students with greater hands-on skill achieved higher online writing scores when holding constant their performance on a paper-and-pencil writing test (Horkay et al., 2006).

Pearson has conducted a number of studies comparing performance of computer-written and handwritten essays for the Texas Assessment of Knowledge and Skills (TAKS). Our findings over time have been mixed in a manner similar to the general research literature. In an initial study, students responding by computer performed lower than a comparison student group who wrote an extended essay by hand (Way, Davis & Fitzpatrick, 2006; Way & Fitzpatrick, 2006). However, a more recent TAKS study indicated that

keyboarded essay performance was higher than written essay performance when students in the computer condition utilized an interface that included standard word processing supports (e.g., cut, copy, paste, undo) and a mechanism to easily gauge how much they had written compared to the total amount of space allotted for the essay response within the system (Davis, Strain-Seymour, Lin, & Kong, 2008).

The emphasis on studying the comparability of computerized and paper-based assessments is motivated by professional guidelines such as the joint *Standards for Educational and Psychological Testing*, which recommend comparability studies for reasons related to the validity of score interpretations (AERA, APA, NCME, 1999). However, as the research findings discussed above illustrate, it is difficult to disentangle effects due to the mode in which writing assessments are administered from the word processing experiences of the participants. Thus, comparability studies seem to be of somewhat limited in informing the transition of a paper-based writing assessment to one that is computer-based. The literature suggests that although the comparability of typewritten and handwritten essays deserves consideration, the more important validity issue may be the need for students to be able to take the statewide writing assessment in a manner that is consistent with how they write in their daily classroom instruction. In this case, if modern word processing tools serve to make students better writers, designing a computer-based writing assessment to achieve comparability with a paper-based writing assessment could serve the wrong purpose, that is, it could negate the educational benefit of the technological advancement.

### **Rater effects in scoring handwritten and computer-based writing assessments**

One consistent finding in the research literature related to the comparability of handwritten and computer-based writing assessments is that human scorers tend to rate typewritten essays more harshly than handwritten essays. A recent study speculated that the most plausible explanation for these findings seems to be that typed essays are more likely to be perceived by scorers as final drafts, and therefore expectations are slightly higher and errors in grammar or spelling are judged more harshly than they are in handwritten essays (Breland, Lee, & Muraki, 2005).

Pearson has implemented specific rater training procedures to offset the tendency of scorers to rate typewritten essays more harshly. These training materials explicitly identify several possible scoring biases typically associated with typed responses:

- Typed essays generally appear shorter than identical handwritten responses. Two pages of handwritten material may take up only a screen or less when the response is typed. Raters must be sensitive to these differences in response length.
- When essays are handwritten, scorers may give writers the benefit of the doubt if spelling or punctuation, for example, is not clear. Because errors may be more easily discerned in typewritten essays, raters must take care not to be unduly influenced by the more visible errors.

- Errors made more evident by typing usually represent only one aspect of the scoring rubric -- that which relates to grammar, usage, and mechanics -- and raters should be certain to evaluate essays based on all criteria on all aspects of the scoring guide.
- Just as the style, tidiness, size, or any other characteristic of student handwriting can inappropriately influence the scores assigned by raters, the greater ease of reading typed responses can affect impressions of student essays. Raters must be constantly reminded to evaluate student responses based on the anchor papers and scoring guides rather than the ease or difficulty of reading responses.

When the Pearson Performance Scoring Center (PSC) evaluates a mix of handwritten and typed essays, the training materials and anchor papers are prepared both in handwritten and typewritten format, and raters alternate between these materials during training until they are comfortable that they are not being influenced by irrelevant differences in the mode of essay entry.

### **The importance of the essay interface**

Still another important aspect of assessing writing by computer concerns the interface that students use to enter their essay responses. Pearson has conducted a number of usability studies that have involved in-depth interviews with individual students to understand what functionality is most necessary to support in an online essay-writing interface (Davis, Strain-Seymour, Lin, & Kong, 2008). Our user interviews revealed some key differences in how students approach essay writing on a computer versus using pencil and paper. When students were asked about their writing method while taking paper-based TAKS writing tests, the responses fell along a continuum from extensive pre-writing to virtually no planning before beginning the essay response. Conscientious pre-writers who wrote drafts on scratch paper before copying their essays into the test booklet reported considerable fatigue. Those students who did not engage in any outlining or rough draft creation expressed dissatisfaction with the limited tools for revising their essays: erasing, crossing out text, and drawing arrows to insert paragraphs. When students compared their writing methodology on the paper-based TAKS test to the writing method that they employ when using a computer to construct an essay, the pre-writing phase was less distinct; outlining, rough drafting, and final writing tended to flow into one another. For many students, some of the time spent on pre-writing shifted to time spent on revising. Over half of the students interviewed reported some significant change in writing process when using pencil and paper versus a computer. And nearly all of those students felt that their computer-based writing method was more similar to their preferred or typical process.

The goal in the interface re-design was to support this diverse range of writing processes and to provide adequate functionality without disadvantaging students who lacked extensive word processing skills. Towards this goal, the decision was made to continue to provide scratch paper to support those students who felt most comfortable with handwriting for certain tasks, including outlining and the creation of graphic organizers. Students' descriptions of essay revising suggested two important areas of functionality within the interface. First, traditional text editing tools – cut, copy, and paste – need to be

designed to align with the practices of self-reported computer experts who tend to almost exclusively use keyboard shortcuts, as well as to support less savvy computer users who may need more visually apparent functionality. The result was the online writing interface presented in Figure 1 with its highly visible tools, support for keyboard shortcuts, and complete interchangeability between these two methods of text editing. The second area of concern to students in planning and revising their essays was space management. While the two-page limit is visually evident while filling in two facing pages in a test booklet, usability testing revealed that a scrolling window was frustrating to students who wanted a concrete sense of the available space remaining. Rather than a character count or a progress bar view of space remaining, the final design included a page view of remaining space, as shown on the right side of Figure 1. Students reported comfort with this display of space remaining, citing both similarities to the paper booklet and to word processing programs' "print preview" functionality.

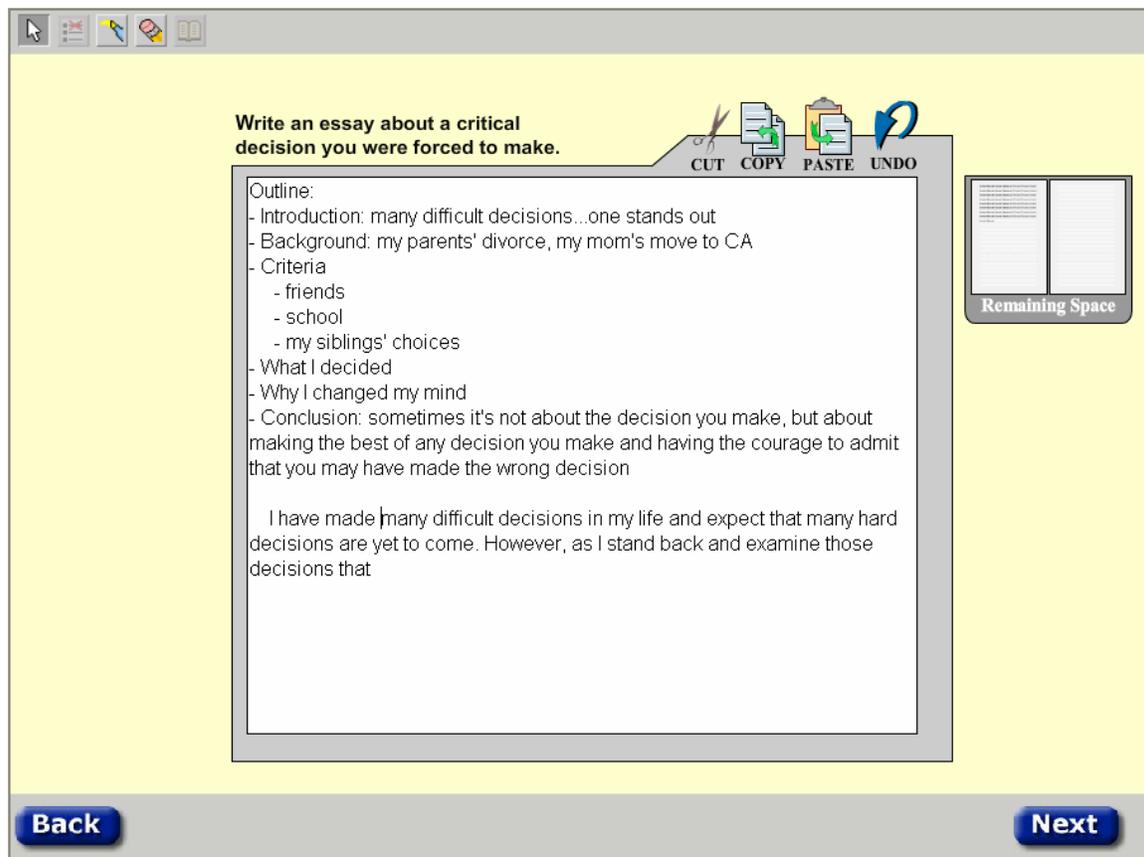


Figure 1. Example of Word Processing Features in Pearson's Online Writing Interface

## **Considerations in Establishing a Statewide Computer-Based Direct Writing Assessment**

In putting together a plan for a statewide computer-based direct writing assessment, there are probably many considerations that should be taken into account. We have identified some specific questions in particular that we think are worthy of discussion:

- Should an online writing assessment be a choice or a requirement?
- Do schools have the infrastructure needed to assess the writing of all students online?
- What other advantages can be gained from assessing writing online?

We discuss these questions in the sections below, and suggest corresponding answers.

### **Should an online writing assessment be a choice or a requirement?**

It is possible to consider introducing writing online as either a choice or a requirement within a statewide writing assessment. The advantage of an optional approach is that schools are able to assess students' writing consistent with the way each student has been instructed and/or the delivery mode each student is most comfortable with. The disadvantage of this approach is the need to support the writing assessment in both modes, along with the associated costs and administrative burdens. In addition, concerns with the comparability of the writing assessment results across the computer-based and handwritten modes may persist as the program is established. These comparability questions are difficult to address, as noted earlier in this paper.

Given the trend towards universal use of word processing software by students for everyday writing, it stands to reason that requiring most students to respond to the statewide writing assessment by computer makes sense. In such a scenario, it would still be possible to offer a paper-and-pencil version of the writing assessment as an accommodation for those individual students that are not able to take the writing assessment by computer or in schools where writing instruction has not yet incorporated the use of word processing tools.

Perhaps the best compromise between these two options would be to implement a required computer-based writing assessment through a one- or two-year phase-in period. During the phase-in period, schools and students would be given a choice of online or paper-based administrations of the writing assessment, after which time the online writing assessment would be required. Given that NAEP is committed to implementing an online writing assessment by 2011 for grades 8 and 12, one strategy a state might consider is to also target a compulsory online writing assessment by 2011.

It is also important to consider how soon children should be exposed to online writing assessment. The 2011 NAEP assessment will continue to assess students at grade 4 using a paper-and-pencil format because of concerns that students may currently lack the necessary keyboarding expertise and experience with using computers on assessments.

States may wish to establish practice that is consistent with NAEP, or delay computer-based writing assessments at lower grades until more definitive evidence about online writing instruction for these students is established.

**Do schools have the infrastructure needed to assess the writing of all students online?**

A barrier in the implementation of online testing is the initial and ongoing investment in computer technology and support services necessary for schools to adequately administer online tests. In general, school districts vary widely in their technological capabilities. An example of this variation is seen in Texas, where schools annually provide information about the extent of their technology infrastructure (Texas Education Agency, 2007). Table 1 summarizes results from 2006-2007 and indicates that over 40 percent of Texas campuses classify themselves in the “Early Technology” or “Developing Technology” categories. Many statewide testing programs encounter similar variations in technology infrastructure across the schools and districts in their states.

Table 1: Texas Campus Statewide Summary of Technological Infrastructure<sup>1</sup>

Technological Infrastructure Classification	Description	#Campuses	%Campuses
Early Technology	There are 10 or more students per computer. There is dial-up connectivity. There is no web-based learning. There is shared use of technology resources.	174	2.2%
Developing Technology	There are 5-9 students per computer. There is direct connectivity to the Internet in 50% of classrooms and library. Most rooms are connected to WAN/LAN. One educator per computer, shared use of other resources.	3044	39.3%
Advanced Technology	There are 4 or less students per computer. There is direct connectivity to Internet in 75% of classrooms and library. Web-based learning is available. All rooms are on LAN/WAN. There is one educator per computer. There is shared use of other resources.	4130	53.3%
Target Technology	There is on-demand access for every student, direct connectivity available in all rooms and web-based resources in multiple rooms. All rooms are connected to WAN. They are fully equipped with appropriate technology.	404	5.2%

<sup>1</sup> It is worth noting that in this same survey, 73.6% of Texas campuses classify themselves as “Developing Technology” with respect to Teaching and Learning. This category is described as follows: “Instruction is teacher-directed and students regularly use technology on an individual basis to access electronic information and develop communication and presentation projects. There is minimal use of technology in foundation TEKS. Most Technology Applications TEKS are met K-8; high school campuses teach at least 2 Technology Applications courses.” (Texas Education Agency, 2007).

For schools with developing technology or even those with advanced technology as defined in Table 1, testing all students by computer for a given assessment can take days or even weeks. Online testing enhances security by eliminating problems with distributing and securing test materials. However, another aspect of test security has to do with the exposure of test questions (or writing prompts) administered online during a testing window that is longer than a traditional paper-and-pencil testing period. For schools lacking sufficient computers to test all of their students quickly, an easily implemented solution is to lengthen the number of days over which testing occurs. That is, it seems reasonable that all schools could give all of the students in a given grade an online writing assessment given a long enough testing window. However, a single writing prompt is likely to quickly become known as it is exposed over time.

One potential solution to this dilemma has been implemented for the Graduate Management Aptitude Test® (GMAT®).<sup>2</sup> This program is administered on demand by computer to potential graduate business school applicants. The exam includes an analytical writing assessment that consists of two 30-minute writing tasks -- Analysis of an Issue and Analysis of an Argument. For each task, the writing prompt is selected at random from a large number of available prompts. The complete set of prompts is made available online for GMAT candidates to review prior to taking the exam. This is not seen as a security risk because the number of prompts is too large for candidates to employ effective rote strategies. The prompts within each type are designed to be as parallel as possible and are scored using the same rubric. Because the prompts are assumed to be parallel, no equating or scaling is applied and final scores are based on the sum of the rater scores obtained from applying the appropriate rubric. It seems very possible that a statewide online writing assessment could be established based on the same or a similar approach, provided that the associated psychometric issues are carefully addressed.

### **What other advantages can be gained from assessing writing online?**

Beyond the validity case for assessing writing by computer which has been the central thesis of this paper, there are two additional advantages to assessing writing online that policy makers may wish to consider. The first advantage is that automated essay scoring (AES) systems that can be applied to evaluate and score typewritten essay responses. The field of AES is relatively new, but has grown and received increased attention within the assessment literature in recent years (Phillips, 2007; Dilkli, 2006).

Commercially-available AES systems are marketed by a number of companies. Pearson's system is known as the Intelligent Essay Assessor (IEA). IEA is a set of web-based software tools for scoring the quality of short to moderate length essays on specific content-oriented or writing ability prompts. IEA uses Latent Semantic Analysis (LSA), a machine-learning algorithm that accurately mimics human understanding of language. This technique can provide instantaneous, individually tailored feedback with scores that

---

<sup>2</sup> For more information, see <http://www.gmac.com/gmac/thegmat/gmatbasics/whatthegmatmeasures.htm>.

have proven to be an accurate measure of the quality of the essays as judged by human readers.

In a statewide computer-based writing assessment, Pearson is able to process and route student responses to the IEA for nearly immediate scoring. One way that IEA (or any AES system) can be utilized in an online writing program is to provide a score that is combined with a single traditional human rating. When the human and computer rating are not the same or adjacent, the essay is routed to a third rater for resolution. This application of AES can potentially cut the time and cost of human scoring by almost half while still providing the reliability of two ratings to each essay response with resolution as needed. This approach is currently being utilized for the online GMAT test (Rudner, Garcia, & Welch, 2006).

A related advantage for assessing student writing online comes through formative writing assessment products that can provide opportunities for students to practice their writing skills throughout the instructional year. These products provide students with immediate, easy-to-understand feedback about their performance and suggestions for how to improve essays and summaries. By applying an AES scoring engine to student responses, the products also relieve teachers from the ongoing burden of grading students' writing. However, the tools also provide teachers with the ability to view class performance summaries, as well as score results and the writing responses of individual students.

A sample output screen from Pearson's formative product, called WriteToLearn™, is shown in Figure 2 and illustrates the typical characteristics of these tools. The information provided to the student includes an overall score and component scores. Additional information includes feedback on essay length, spelling, grammar, and redundancy of information. The product also provides opportunities for students to incorporate the feedback in revising essays that can be resubmitted to the system for scoring.

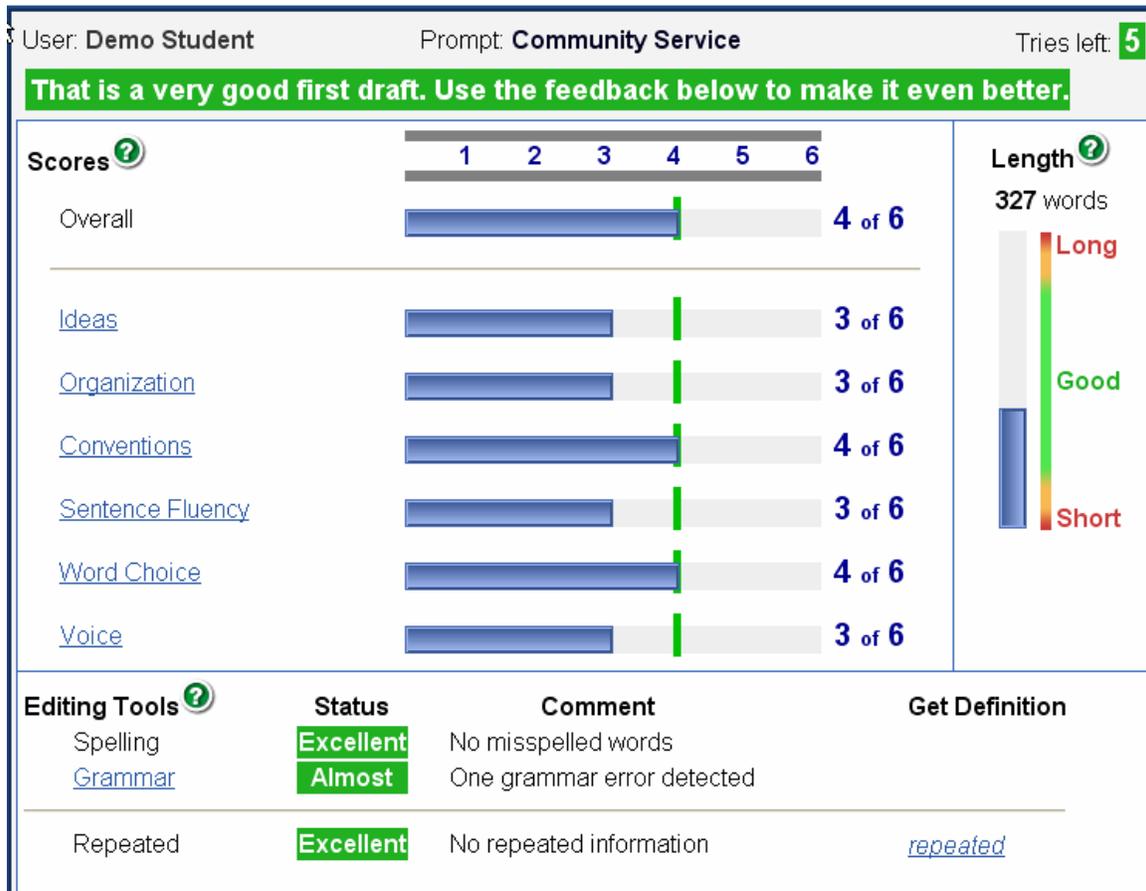


Figure 2. Sample Output from Pearson’s WritetoLearn Formative Writing Product

The success of computer-based formative writing products such as WriteToLearn in effectively supplementing classroom writing instruction suggests that the use of these tools will continue to increase in the future. These tools also underscore the point that statewide writing assessments must inevitably move online to match the way in which students learn and practice writing in the classroom.

### Summary and Conclusions

The purpose of this paper was to present the validity case for assessing writing by computer. This case is supported by survey findings suggesting that most students are using computers and word processing software when they write in academic settings, for many if not all of their assignments. The trend suggested by the current research points to a need for state policy to seriously consider the use of computers and word processing software to measure the writing ability of students in their statewide assessments. In making this validity case, we have illustrated issues and challenges with assessing writing by computer, as well as some questions for policy makers to consider as they address transitioning writing assessments to a computer-based format.

Change in the assessment world is often slow, and accomplishing change in state testing programs is particularly challenging because of competing stakeholder interests and a

preference for minimizing the disruptions that change tends to bring. Sometimes, changes in statewide assessments are made in order to reform instruction, such as adding an essay component to a test to encourage the schools to teach writing. In the case of assessing writing online, however, the impetus for change is coming from the classrooms and the students themselves, and the time to respond is now.

### **About the Authors**

Dr. Walter (Denny) Way is Vice President of Psychometric and Research Services in Pearson's Assessment and Information Group. Dr. Way directs all psychometric and research activities for Assessment & Information and consults with management and clients on psychometric issues. Dr. Way is a nationally known expert in computer-based testing and psychometrics with over 20 years of assessment experience. Prior to working at Pearson, he spent 16 years at Educational Testing Service in Princeton, New Jersey, providing psychometric and management oversight for units serving ETS programs in admissions, higher education, and certification and licensure.

Dr. Laurie L. Davis holds a Ph.D. in Educational Psychology with a specialization in Quantitative Methods and Psychometrics from the University of Texas. She serves on the editorial review board for Educational and Psychological Measurement and is a member of the AERA Division-D mentorship committee. She has published in the area of computerized adaptive testing (CAT) and has presented on comparability methods for computer-based testing at several national conferences. Currently, she works as Manager of Psychometric Services within the Assessment and Information group at Pearson, with primary support to Texas assessment program.

Dr. Ellen Strain-Seymour conducts usability research and works on online testing initiatives within Pearson's Assessment and Information group. She received her doctorate at the University of Southern California and taught human-computer interface and instructional design for eight years as a member of the faculty at Georgia Tech. She has founded and managed a Flash development firm specializing in educational/training applications. Prior to her current position at Pearson, she managed online products and user-centered design initiatives at Pearson Learning and at Harcourt. Dr. Strain-Seymour has published one book and several articles on interface design, media studies, and ethnographic methodologies.

## References

- ACT, Inc. (2007). *Writing framework for the 2011 National Assessment of Educational Progress: Pre-Publication Edition*. Prepared for the National Assessment Governing Board in support of Contract No. ED-05-R-0022. Iowa City, IA: ACT, Inc.
- American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Belden Russonello & Stewart. 2007. *The 2007 Survey on Teaching Writing: American Public Opinion on the Importance of Writing in Schools*. (Conducted for the National Writing Project.) Washington, DC: Belden Russonello & Stewart.
- Breland, H., Lee, Y.W., & Muraki, E.(2005). Comparability of TOEFL CBT essay prompts: response-mode analyses. *Educational and Psychological Measurement*, 65 (4), 577-595.
- Bridgeman, B., & Cooper, P. (1998). *Comparability of scores on word-processed and handwritten essays on the Graduate Management Admission Test*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Davis, L.L, Strain-Seymour, E., Lin, C., & Kong, X. (2008). *Evaluating the comparability between online and paper assessments of essay writing in the Texas Assessment of Knowledge and Skills*. Presentation given at the Annual Conference of the Association of Test Publishers, Dallas, TX.
- Dikli, S. (2006). An overview of automated scoring of essays. *Journal of Technology, Learning, and Assessment*, 5(1). Retrieved December 13, 2007 from <http://www.jtla.org>.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools – A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Horkay, N., Bennett, R.E., Allen, N.L., Kaplan, B., & Yan, F. (2006). Does it matter if I take my writing test on computer? An empirical Study of mode effects in NAEP. *Journal of Technology, Learning, and Assessment*, 5(2). Retrieved December 11, 2007 from <http://www.jtla.org>.
- Phillips, S.M. (2007). *Automated essay scoring : A literature review*. (SAEE research series #30). Kelowna, BC: Society for the Advancement of Excellence in Education.

- Piccaino, A.G., & Seaman, J. (2007). *K-12 online learning: A survey of U.S. school district administrators*. Needham, MA: The Sloan Consortium. Available at: [www.sloan-c.org-publications-survey-pdf-K-12\\_Online\\_learning.pdf](http://www.sloan-c.org-publications-survey-pdf-K-12_Online_learning.pdf).
- Rudner, L. M., Garcia, V., & Welch, C. (2006). An Evaluation of the IntelliMetric<sup>SM</sup> Essay Scoring System. *Journal of Technology, Learning, and Assessment*, 4(4). Retrieved December 13, 2007 from <http://www.jtla.org>.
- Russell, M., & Plati, T. (2001). Effects of computer versus paper administration of a state-mandated writing assessment. *TCRecord*. Retrieved December 12, 2007, from <http://www.tcrecord.org/Content.asp?ContentID=10709>.
- Russell, M., & Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, 5(3). Retrieved December 12, 2007, from <http://epaa.asu.edu/epaa/v5n3.html>.
- Solomon, C., Lutkus, A. D., Kaplan, B., Skolnik, I. (April 2004). *Writing in the nation's classrooms: teacher interviews and student work collected from participants in the 1998 NAEP writing assessment* (ETS NAEP Technical and Research Report). Princeton, NJ: Educational Testing Service.
- Texas Education Agency (2007). *Campus statewide summary by key area*. Retrieved December 13, 2007 from: <http://starchart.esc12.net/docs/campusSWSummarybyKA20062007.pdf>
- Way, W.D., Davis, L.L., & Fitzpatrick, S.J. (2006). *Score comparability of online and paper administrations of the Texas Assessment of Knowledge and Skills*. Paper presented at the annual meeting of National Council on Measurement in Education, San Francisco, CA.
- Way, W. D., & Fitzpatrick, S. (2006). *Essay responses in online and paper administrations of the Texas Assessment of Knowledge and Skills*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.