Computers & Education 101 (2016) 115–131

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The effects of prior computer use on computer-based writing: The 2011 NAEP writing assessment

Tamara P. Tate [a, *], Mark Warschauer [a], Jamal Abedi [b]

[a] School of Education, University of California, 3200 Education, Irvine, CA 92697-5500, United States
[b] School of Education, University of California, One Shields Ave., Davis, CA 95616, United States

A R T I C L E I N F O

Article history:
Received 1 October 2015
Received in revised form 19 April 2016
Accepted 2 June 2016
Available online 7 June 2016

Keywords:
Computer-mediated communication
Secondary education
Improving classroom teaching
Pedagogical issues
Teaching/learning strategies

A B S T R A C T

Writing achievement levels are chronically low for K-12 students. As assessments follow the transition to computer-based writing, differences in technology access may exacerbate students' difficulties. Indeed, the writing process is shaped by the tools we use and computer-based writing is different from writing with pen and paper. We examine the relationship between reported prior use of computers and students' achievement on the first national computer-based writing assessment in the United States, the 2011 National Assessment of Educational Progress (NAEP) assessment. Using data from over 24,100 eighth grade students, we found that prior use of computers for school-related writing had a direct effect on writing achievement scores on the computer-based NAEP assessment. One standard deviation increase in prior use led to a 0.14 and 0.16 standard deviation increase in mean and scaled writing achievement scores respectively, with demographic controls and jackknife weighting in our SEM analysis. We also looked at earlier NAEP assessments and found that prior computer use did not positively affect the earlier pen and paper-based writing assessments.

1. Introduction

Writing is a complex and highly challenging activity (Deane, 2011). It is not only a problem-solving process, but also a constructive process of transforming, formulating, and constituting new knowledge (Bazerman, 2011). Most learners struggle with the prerequisite coordination of multiple processes and linguistic conventions (DeBono, Hosseini, Cairo, Ghelani, Tannock, & Toplak, 2012; De La Paz & Graham, 2002; Deane et al., 2008). For decades, the National Assessment of Educational Progress (NAEP) has tested U.S. students in a number of disciplines, including writing. NAEP has shown that the majority of students are not even minimally proficient writers, let alone skillful ones, with only 27 percent of all students, 11 percent of Black students, and 14 percent of Hispanic students at or above proficient levels (NCES, 2012). Similarly, the College Board (2015) has announced that the SAT writing results continue to decline at a rate nearly twice as large as the declines in math and reading over the same period. In addition, despite its importance and complexity, writing receives less instructional attention than subjects like reading and math, particularly in the elementary and middle school grades (Lyon & Weiser, 2013; Warschauer, 2011; Graham & Perin, 2007). Nonetheless, writing is connected to all content areas and the deficiencies in

* Corresponding author.
E-mail addresses: tatet@uci.edu (T.P. Tate), markw@uci.edu (M. Warschauer), jabedi@ucdavis.edu (J. Abedi).

http://dx.doi.org/10.1016/j.compedu.2016.06.001
0360-1315/

students' writing proficiency are hindering their development of academic English (Zheng & Warschauer, 2015) and subsequent college and career readiness (Graham & Perin, 2007).

Our society calls for vastly complex and ever-changing genres and text modalities to be learned. Children should be prepared for these evolving practices; in fact the Institute for Education Sciences (IES) Practice Guide recommends that students be taught to use the writing process for a variety of purposes and become fluent in multiple modalities of transcription including word processing. In particular, today's students need to successfully negotiate computer-based writing in order to have equal access to college and career options (cf., Applebee, 2007). High-stakes assessments are migrating to computer-based formats (e.g., Smarter Balanced and PARCC assessments of Common Core State Standards), and gateway tests for higher education are increasingly computer-based. In order for students to emerge from K-12 education "college and career ready"—the goal under the current Common Core State Standards—they need to be able to write using computers. Teaching students current forms of literacy, such as computer-based writing, are important to prepare them to participate fully in the community (Langer, 1991). In many instances, however, students receive inadequate explicit instruction in writing on computers.

These new technologies present new cognitive challenges and opportunities (Bazerman, 2011) that students and teachers will need to address. We know that the writing process is shaped by the author's tools (see discussion in Wertsch, 1991). Each development in technology affects the writing process itself. For example, current research finds that students write more and write better on computers (see discussion in Morphy & Graham, 2012; Collins, Hwang, Zheng, & Warschauer, 2014; Graham & Perin, 2007; Sandene et al., 2005; Russell & Haney, 1997; Russell & Plati, 2002; Applebee & Langer, 2009). This leads us to query how the introduction of a powerful tool such as a computer may transform the writing process and how that transformation may be shaped by prior experiences in individual students' lives.

In order to test the computer-based writing skills of our youth, computer-based writing assessments provide the closest measure (NAGB, 2010). However, most studies of computer writing by and computer assessment of K-12 students have used fairly small samples (see discussion in Bangert-Drowns, 1993). This secondary data analysis looks at the relationship between prior use of computers for writing and achievement on the 2011 NAEP computer-based writing assessment. Our research questions were as follows:

1. Does the prior use of computers positively affect students' results on a computer-based assessment?
   a. Does it matter whether the prior computer use is school-related or personal?
   b. Are reports of school-related use by students or teachers more predictive of improved writing achievement?
   c. Does a teacher's use of technology for writing instruction predict students' improved writing achievement?
   d. Does technology-related professional development for the teacher predict students' improved writing achievement?
2. Does the effect of prior use of computers on writing achievement vary by demographic group?

By understanding the model of how prior use of computers and writing achievement on a computer-based writing assessment relate, we hope to inform both assessment and instructional efforts to teach all students how to write effectively on computers.

2. Conceptual framework

Our work is based on a broad notion of the role of tools, which encompass the mental, linguistic, and physical devices used to enhance writers' performance (Englert, Mariage, & Dunsmore, 2006). We believe that writing is culturally situated and mediated by these tools (Deane et al., 2008; Wertsch, 1998). New technologies allow us to produce, transmit, store, and process written texts (Bazerman, 2011). Each development in technology affects the writing process itself (cf., Berninger & Winn, 2006). For example, some tools may constrain idea generation and elaboration (Berninger & Winn, 2006). Success with composing on these new devices depends upon a willingness and ability to change modes, adapt prior strategies (Cochran-Smith, 1991), and navigate the specific tool affordances that both promote and inhibit good writing. These concerns led us to our research questions, a desire to understand whether (and for who) the prior use of computers (the tool) improves students' writing in a computer-based writing assessment.

We expected that practice using a specific tool would affect the writing process with that same tool. We thought that it was possible for computer use beyond writing for school, such as e-mailing, could provide a comfort level and familiarity with the mode of digital writing that would impact the writing process in an assessment setting. Thus, we initially looked at a wide range of variables related to digital technology use.

Our variable selection was also impacted by our belief that literacy is culturally situated. Because of this, cognitive apprenticeships are important in the acquisition of writing skills. Cognitive apprenticeships teach novices the practices of the community, including the acquisition of the discourse, tools and actions. Teachers can make these practices of the writing process visible; and effective teachers model and describe the knowledge they have about writing (Englert et al., 2006; Vygotsky, 1981). These teachers provide support as novices acquire the discourses, strategies, tools, and actions needed. For this reason, one group of the survey questions examined for our prior use latent variable related to the use of technology by teachers when teaching writing. We also included teacher professional development in technology as a potential component of relevant prior computer use, hypothesizing that increased professional development could lead to increased or improved modeling and direct teaching of the use of technology for writing.

The comprehensive data available from the NAEP 2011 assessment and current statistical methods allow us to look closely at students' computer-based writing. Insights for teaching diverse students to be better writers on computers may arise from better understanding the model of how prior computer use and computer-based writing achievement relate. We are also mindful of a prior study done in preparation for NAEP's implementation of computer-based assessments. In the earlier work researchers compared scores for paper and computer versions of the 2002 NAEP writing assessment and found no significant population-level differences (Horkay, Bennett, Allen, Kaplan, & Yan, 2006; Sandene et al., 2005). A repeated-measures analysis of variance failed to detect a significant effect for delivery mode on achievement score (Sandene et al., 2005), even when controlling for gender, race/ethnicity, parents' education level, eligibility for free/reduced-price school lunch, and school type. Analysis of essay length showed no measurable differences on the number of words written on paper or on computer (Sandene et al., 2005). Sandene et al. (2005) found no equity-related differences between pencil and paper assessments and computer-based assessment at a population (versus individual) level, except with respect to urban students who performed 15 percent higher on paper and pencil tests. There was a small but significant gender effect on writing length but not on scores, with males writing approximately two percent fewer words on paper than on computer for the second task, and two percent more females responding to the second question on paper (Sandene et al., 2005).

Although the 2002 study showed no significant mean score differences between those taking the computer tests and those taking the tests with pencil and paper at a population level, it indicated that computer familiarity did significantly predict performance at an *individual student level* (Horkay et al., 2006). Using students' self-reported computer experience to create a composite score to measure familiarity, Sandene et al. (2005) found no significant effect for prior use of computers on achievement scores, but there was an effect for keyboarding skills. Sample size limitations prevented further analysis of these differences, except for gender, which was inconclusive.

Our study explores the issue of computer familiarity in further depth and at scale across a nationwide sample nearly a decade later. With the growing prevalence of computer-based writing and writing assessment, concerns about the impact on students without regular access to digital technology abound. There are widely differing degrees of school and home computer access and use; research shows that social factors can be even more important than technical factors in shaping productive use of technology (Warschauer, 2011). Further, demographic impacts may exacerbate differences among students' computer skills in ways that need to be understood and, perhaps, addressed. Large-scale assessments like NAEP are useful for exploring these types of achievement gaps and differences among groups (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007).

3. Methods

3.1. Data source

This research analyzed the data from over 24,600 eighth grade students. NAEP assessments are widely regarded as high quality, with strong construct and measurement validity (see, e.g., Wenglinsky, 2005). The weighted national school participation rates for the assessment were 97 percent (100 percent for public schools; National Center for Education Statistics (NCES) 2012). To the extent certain subgroups fell below 70 percent, NCES conducted an analysis of potential bias. Compared with the distribution of all eligible students, the distribution of the weighted sample did not differ with respect to any of the variables utilized in this analysis (Rogers, Stoeckel, & Sikali, 2013). This analysis utilized the restricted data set, which includes scaled and raw scores, detailed survey data, and individual keystroke data. As suggested by NCES, the following were treated as missing: multiple responses, responses not reached or administered, omitted responses, non-ratable responses, illegible responses, and off task responses.

3.1.1. Sampling

NAEP sampling techniques strive to create a representative nationwide sample of students in grades 4, 8, and 12. Participants are selected using a stratified cluster sampling, where the population is divided into different strata, or geographic areas of interest, from which the schools were selected (Beaton et al., 2011). In order to approximate the population, sample weights are used to correct for oversampling of certain low incidence populations and adjust the overall results by the actual population proportion (Johnson, 1992). These weights allow for valid inferences to be made about the population (Beaton et al., 2011).

While traditional analysis procedures assume that observed data from different individuals are independent of each other and randomly distributed, NAEP results may be stratified and clustered (Johnson, 1992; Zwick, 1987). Through clustering, weighting, and marginal estimation procedures, NAEP allows for population and group estimates (Beaton & Zwick, 1992). Ignoring these effects leads to biased estimates of variance and generally to underestimating the biases (Johnson, 1992). In addition, the deeply stratified cluster samples influence the likelihood ratio and inflate the differences in the chi-squares, and the design effect for item p values is estimated to be roughly 2 (an estimate of percent of examinees with a given response pattern should be equivalent in precision to a simple random sample approximately half as large; Haertel, 1984). Where our current analysis looked at the individual booklet-level responses, weighting was not applied (Allen & Donoghue, 1996). Where we looked at aggregated values of individual, unscaled scores, we used jackknife weights for the analysis.

3.1.2. Population

We focus on eighth grade students, as prior research suggests that the middle school years are critical for the development of academic writing (De La Paz & Graham, 2002; Zheng & Warschauer, 2015). Indeed, some refer to an eighth-grade literacy cliff (e.g., Zheng & Warschauer, 2015). In addition, the NAEP data for eighth grade, unlike twelfth, has a teacher survey reporting on computer use allowing for potential correlation between student-reported use and teacher-reported use.

3.1.3. Assessment

NAEP assessments are known for their robust construct validity (see, e.g., Applebee, 2007; Wenglinsky, 2005). Over the course of 18 months, more than 500 individuals developed a framework for the NAEP writing test (NAGB, 2010). These individuals included leading educators and experts in the field of assessment (Applebee, 2007). The NAEP writing framework was designed to reflect the way students write today, using word processing software and commonly available tools (NAGB, 2010). As a result, the assessment allowed students to use common word-processing tools:

- editing (cut, copy, paste, delete, backspace);
- formatting (indenting, underlining, highlighting, bolding, and italicizing);
- spelling check, grammar check, thesaurus, and dictionary; and
- viewing and reviewing during the assessment (NAGB, 2010).

In addition, the assessment included student and teacher surveys that gathered information about students' and teachers' prior use of computers.

The students were given two different writing tasks and had 30 min to complete each task. There were a total of 22 writing prompts in three areas: (a) to persuade, (b) to explain, and (c) to convey experience, either real or imagined. Responses were scored by three trained evaluators on a 6-point scale—effective skill, adequate skill, developing skill, marginal skill, and little or no skill—across three areas of writing: (a) development of ideas, (b) organization of ideas, and (c) language facility and conventions (NCES, 2012; NAGB, 2010). NAEP evaluators used holistic scoring rubrics to evaluate the response as a whole, rather than assessing independent parts of the response (NAGB, 2010). NAEP ensures scorers' reliability through back reading where scoring supervisors selectively review at least five percent of each scorer's work, periodic calibration of multiple scorers, and an inter-rater reliability statistics check of 25 percent of the responses (NCES, 2009).

The presentation of the items was alternated so that the same item appeared first in some booklets and second in others. This balancing of order of presentation is important, because NAEP has found that assessments with open-ended responses show decreased scores in the later items (Johnson, 1992). However, because of this balanced incomplete block (BIB) spiraling method of sampling, NAEP data is complex to model (Beaton & Zwick, 1992). BIB spiraling presents each item to a large number of participants and pairs items to allow correlations between the items (Beaton et al., 2011; Johnson, 1992). BIB spiraling is used to allow broad coverage of the item pool, yet not impose excessive test taking upon individual students (Beaton et al., 2011; Applebee, 2007). This design has two major limitations. Each student only receives a fraction of the test, which weakens any ability to determine individual achievement in any of the subject areas (and increases measurement error). Second, the scores from each block may not be highly comparable (Beaton et al., 2011; Applebee, 2007).

3.2. Variables

Initial variable selection included: (a) reported prior computer use; (b) achievement scores on the writing assessment, either scaled or mean scores; and (c) group variables for gender, ethnicity/race, eligibility for free/reduced lunch, highest level of parental education, English-learner status, and disability status.

3.2.1. Achievement measure

NAEP uses Item Response Theory (IRT) scaling to allow for estimates of item characteristics and difficulties, as well as multiple imputation to estimate student achievement values in terms of plausible values (Beaton et al., 2011; Zwick, 1987). However, the NAEP Primer cautions that researchers interested in interaction effects should not work with plausible values, but should perform their own marginal estimation (Beaton et al., 2011). Thus, this analysis looked at both the mean of all the individual raters' scores for a particular student's response and scaled scores at the booklet level and as an aggregated group, but did not use plausible values since we were interested in the possible interactions between prior use and our demographic variables.

The scaled booklet-level scores (−2.18 to 3.04) were used as the achievement variable or dependent variable for the initial analyses. This allowed us to get a sense of student performance normed (using IRT) across the various booklets and allowed analysis of a larger numbers of cases for low incidence demographics. The main purpose of NAEP's IRT analyses is to provide a common scale on which to compare achievement across groups (Messick, Beaton, & Lord, 1983). Researchers can then compare performance across groups, if the subgroups are of sufficient size (Messick et al., 1983).

Additional analysis of student scores was done with the mean of the unscaled scores (interval scale, 1–6) sorting the students and analyzing them by booklet. However, the booklet grouping reduces the number of students in the smaller incidence groups, such as students with disabilities. The final analysis used the mean score for each individual participant and considered aggregate effects.

The scaled writing achievement score had a mean of −0.04 and a standard deviation of 0.96. The raw scores had a mean of 2.64 and a standard deviation of 0.98 (see Table 1, for additional descriptive statistics on the scaled and mean score variables). Both outcome variables were quite close to a standard curve, with a slight skew (particularly for the mean scores) to the right. Thus, our use of linear regression is supported by these descriptive statistics.

3.2.2. Prior computer use and access

The NAEP data includes survey information from teachers and students with respect to two primary background variables, computer use (especially with respect to writing with computers) and types and amount of writing practice. This research focused on the first background variable, amount of prior computer use, using the responses to questions related to prior use and access to create a latent construct (Haertel, 1984).

Variables relating to prior computer use and access included student-reported measures of how often: (a) the Internet is used to get information, (b) a computer is used for a first draft, (c) a computer is used to make changes in writing, (d) a computer is used to complete writing, (e) a computer is used to write school assignments, (f) a computer is used to write not for school, (g) a computer is used for emails, and (h) a computer is used to write on the Internet. Similar teacher-reported measures of students' classroom use of computers for writing, teacher use of technology in the classroom, and teacher professional development relating to technology use were used. We also considered the effect of having a computer at home, but over 90 percent of the students reported having a computer at home so the variable was of little predictive value. The prior use variables were all standardized for the analyses.

3.2.3. Group variables

We looked at differences in various demographic groups: (a) gender, (b) national school lunch eligibility and parental education (as proxies to indicate socioeconomic status), (c) English language learner status (prior, current, or not applicable), (d) students with individualized education plans (IEPs) or 504 plans under the American with Disabilities Act, and (e) race/ethnicity. Dichotomous variables were created for these groups. We considered both parental high school completion and parental college completion as potential control variables. Ultimately, we used the parent college completion variable; literature suggests that first generation college students have unique challenges and that as such it is a useful designation for understanding certain aspects of socioeconomic status (see, e.g., Bowen, Kurzweil, & Tobin, 2005; Saenz, Hurtado, Barrera, Wolf, & Yeung, 2007; Sirin, 2005; Pascarella & Terenzini, 1991; Jackman & Jackman, 1983; Snibbe & Markus, 2005). In addition, the high school completion variable had less predictive strength, because only 9 percent of the students had parents who did not complete high school, whereas 44 percent had parents who did not complete college.

3.2.4. Variables used in quasi-longitudinal analysis

The 2007 assessment used in our quasi-longitudinal analysis had several differences from the 2011 assessment beyond the mode switch from paper to computer in 2011. The changes include somewhat different frameworks, with slightly adjusted emphases on genres, a decision in 2011 to include a specifically designated audience for the writing task, and different questions on the student survey. Although we note these differences, we believe that a comparison of the 2007 and 2011 assessments for the purpose of determining the impact of prior computer use on writing achievement in different modes is

Table 1
Standardized coefficients and z scores for latent variables in final structural equation model, using scaled and mean writing dependent variables, with controls and jackknife weighting.

Observed variable	Latent construct	Scaled β	Mean β
Draft/revise	Teacher-reported	0.79 (0.02) z 39.51	0.79 (0.03) z 30.09
Complete	Teacher-reported	0.85 (0.01) z 61.89	0.85 (0.02) z 44.76
Word processing	Teacher-reported	0.78 (0.02) z 45.92	0.78 (0.02) z 36.66
Use internet	Teacher-reported	0.63 (0.03) z 21.63	0.63 (0.04) z 17.74
Use Internet	Student-reported	0.73 (0.01) z 71.15	0.73 (0.01) z 70.18
First draft	Student-reported	0.74 (0.01) z 59.89	0.74 (0.01) z 60.747
Make changes	Student-reported	0.85 (0.01) z 155.97	0.85 (0.01) z 158.64
Complete	Student-reported	0.79 (0.01) z 130.35	0.79 (0.01) z 128.34
Write for school	Student-reported	0.59 (0.02) z 37.11	0.59 (0.02) z 37.22

Note: Standard errors in parentheses.

illustrative and consistent with our analysis of the 2011 results. The scaled student scores (scaled −2.18 to 3.04; mean −0.04; sd 0.96 in 2011; scaled −2.30 to 3.10; mean −0.04; sd 0.96 in 2007) were used as the initial achievement variable in the analyses.

Our analysis of the 2011 assessment suggested that student reports of computer usage were more predictive of writing achievement than teacher reports of classroom computer usage. In addition, it suggested that school-based use, rather than recreational use, of computers was more associated with achievement levels on the assessment. Therefore, we chose as our independent variable from the 2011 survey "how often do you use a computer to write school assignments" and from the 2007 survey, "write paper for school - use computer from beginning." We refer to these measures as "prior use."

The 2007 survey scale was 1 (never), 2 (sometimes), and 3 (almost always); the 2011 scale was 1 (never or hardly ever), 2 (once/twice a month), 3 (once or twice a week), and 4 (every day or almost). We considered 1 on both scales to be the same, determined that once or twice a week was close to "sometimes" and coded both 2, and combined the 3 and 4 values into a single 3 value, so that weekly and daily use were both coded the same as "always or almost always." We checked our calculations by combining the 2 and 3 values instead, and found the same trends and levels of significance, with generally decreased coefficients on the prior use variable. Nonetheless, we believe that "weekly" and "daily" are closer to "almost always" than "monthly" and "weekly" are to "sometimes" and have presented our results in a consistent manner.

3.3. Analytic methods

We used Stata Version 14.0 SE statistical software to analyze the results of the 2011 NAEP writing assessment.

3.3.1. Structural equation modeling

The analysis included structural equation modeling (SEM) of the data using both the IRT scaled scores ("scaled scores") and the mean of the individual scores by trained reviewers on each essay ("mean scores") at an aggregate (all essays, regardless of different writing tasks) and booklet-level analysis (isolating each writing) to check for robustness and comparability. Analysis separating the criterion instrument into booklets addresses the BIB spiraling in NAEP instruments (Haertel, 1984; Welch, Anderson, & Harris, 1982). Thus, in order to observe individual-level correlations, this study looked at each booklet individually, then cross-validated the results among booklets (see Allen & Donoghue, 1996). All booklets are considered parallel in BIB sampling, so the expectation was that the booklet-level results would not be statistically different from aggregate results, and the missing data can be regarded as random (Zwick, 1987).

The use of SEM allowed us to model the potential causal relationships between prior computer use and achievement scores (cf., Schreiber, Nora, Stage, Barlow, & King, 2006). SEM allows simultaneous estimation of the full model parameters and offers flexibility in modeling reciprocal relationships and creating latent variables, greatly enhancing the ability to analyze the NAEP data (Abedi, 2002; Messick et al., 1983).

The NAEP assessment surveys students and teachers about a number of items including the amount of time they spend on certain types of computer-based writing and related tasks. Based on our theoretical belief that the tools used in writing affect the writing process and that teachers modeling of writing with technology could be an important factor, we chose a total of 28 survey questions related to prior technology use and access by both students and teachers, for our initial analysis. These questions included: (a) student reports of frequency of computer-based writing for school purposes, (b) student reports of frequency of computer-based writing for outside of school uses, (c) teacher reports of school-related computer-based writing, (d) teacher reports of the use of technology for writing instruction, and (e) teacher reports of professional development related to technology and instruction. We sorted the 28 variables into 5 latent variables reflecting these categories. We correlated teacher and student reports of school use of computers for writing, because both teachers and students should both be reporting at least a portion of the same school writing assignments using computers. We also correlated teacher instruction with digital devices and teacher professional development related to digital technology, because teachers with more professional development in the use of technology should be more comfortable and therefore inclined to use devices in instruction. Next, we controlled for student gender, ethnicity, ELL status, disability, free and reduced lunch status, and whether or not a parent graduated college. Estimations were done using maximum likelihood with missing variables.

Our initial SEM contained a number of variables with factor loadings below 0.40 or that were statistically insignificant. Based on the direct effects and factor loadings found in our SEM results, we created a more parsimonious model. We removed any questions from the latent variable that was not significant. Our final model retained both teacher and student-reported use of computers for school-related writing (see Fig. 1). Finally, jackknife weighting was used (sampling units, PSUID; Strata, REPGRP1; Sample weight, ORIGWT; Student Replicate Weights, SRWT01-62).

Our booklet level analysis used Stata's "if" function to generate direct effects for each latent variable on the mean individual writing scores for each of the 22 writing tasks. The same controls as in the aggregate analysis were used.

3.3.2. Regression

As a robustness check, we also used ordinary least squares (OLS) regression to look at the relationship between reported prior computer use and achievement scores. We tested the following model: *Writing Achievement* $= B_0 + B_1$*(teacher-reported prior use)* $+ B_2$*(student-reported prior use)* $+ B_3$*(female)* $+ B_4$ *(Black)* $+ B_5$ *(Hispanic)* $+ B_6$ *(Asian)* $+ B_7$ *(other)* $+ B_8$ *(free/reduced lunch)* $+ B_9$ *(college graduate parent)* $+ B_{10}$ *(former ELL)* $+ B_{11}$ *(current ELL)* $+ B_{12}$ *(student with disability)* $+ B_{13}$ *(teacher-reported prior use)(female)* $+ B_{14}$ *(teacher-reported prior use) (Black)* $+ B_{15}$ *(teacher-reported prior use) (Hispanic)* $+ B_{16}$ *(teacher-reported*

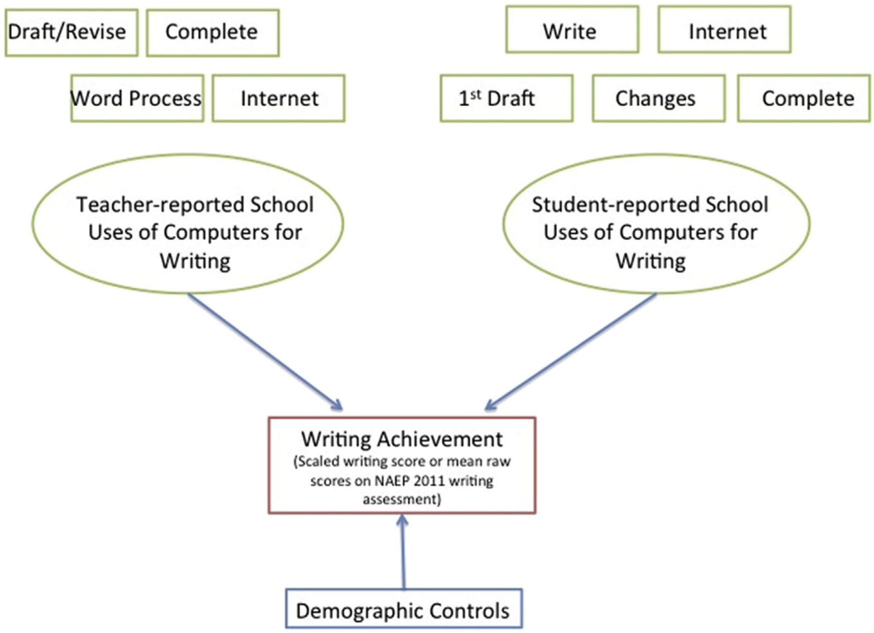

Fig. 1. Parsimonious final structural equation model showing direct effects of latent variables on writing score.

prior use) (Asian) + B_{17} *(teacher-reported prior use) (other)* + B_{18} *(teacher-reported prior use) (free/reduced lunch)* + B_{19} *(teacher-reported prior use) (college graduate parent)* + B_{20} *(teacher-reported prior use) (former ELL)* + B_{21} *(teacher-reported prior use) (current ELL)* + B_{22} *(teacher-reported prior use) (student with disability)* + B_{23} *(student-reported prior use)(female)* + B_{24} *(student-reported prior use)(Black)* + B_{25} *(student-reported prior use)(Hispanic)* + B_{26} *(student-reported prior use)(Asian)* + B_{27} *(student-reported prior use)(other)* + B_{28} *(student-reported prior use)(free/reduced lunch)* + B_{29} *(student-reported prior use)(college graduate parent)* + B_{30} *(student-reported prior use)(former ELL)* + B_{31} *(student-reported prior use)(current ELL)* + B_{32} *(student-reported prior use)(student with disability)* + *e*. We used the standardized coefficients from the final SEM to create the weighted teacher-reported and student-reported prior use variables (Table 1). Regressions were done using each of the scaled writing score and the mean writing score to operationalize writing achievement. Linear regression was appropriate for our data, which showed little skewness or kurtosis.

3.3.3. *Factor analysis*

As a second robustness check, we used factor analysis to confirm our latent variable construction. Stata's principal factor analysis attempts to identify a small number of latent variables or dimensions that explain the shared variance of a set of measures. We used the Kaiser criterion (Kaiser, 1960), initially retaining factors with eigenvalues greater than 1 as our lower bound, but also considering whether that bound was appropriate for our data by looking at scree plots for large drops in values. If an item has a loading of over 0.40 on a factor, it may be considered a good indicator of that factor, although there is debate on exact levels. Finally, we considered rotations to see if they would improve our analysis, particularly an oblique rotation, which allowed for correlations between the latent variables as seen in our SEM.

3.3.4. *Quasi-longitudinal analysis*

Finally, we used OLS regression to assess the relationship between achievement scores and prior computer use for an earlier paper-based NAEP assessment in 2007. If prior computer use is, in fact, important for computer-based writing in a unique way, we would not expect to see the same correlation of writing achievement and prior computer use on paper and pen based assessments. As discussed under "Variables," above, there are differences worth noting between the assessments in 2007 and 2011 beyond the fact that 2007 and earlier were paper-based.

4. Results

Our first research questions was, "Does the prior use of computers positively affect students' results on a computer-based assessment?" Writing achievement was measured by the score received by the individual for holistic writing quality (either the scaled score or the mean of the ratings). Prior use was measured by our latent variables based on responses to teacher and student surveys about technology use for school-related writing. By looking at both teacher and student reports, we examined sub-question "a" regarding the predictive value of each. Our initial analyses also included personal uses of computers for writing in order to look at sub-question "b" regarding the impact of both types of use. Finally, our initial analyses included reports by teachers about their use of technology during writing instruction (sub-question "c") and prior relevant professional development in order to assess the affect of those factors on subsequent writing achievement (sub-question "d").

Group effects and interactions were included through the use of individual dummy variables (e.g., female, Asian) for each demographic group, which acted as controls with the prior use variables. We also looked for potential interactions between prior use and our control variables, which allowed us to address our second research question regarding heterogeneous effects.

4.1. Structural equation modeling

We found that student-reported not for school writing, teacher professional development, and teacher instruction using technology were not statistically significant in our SEM of the scaled scores (see Fig. 1 and Table 2). Both student-reported and teacher-reported use of computers for school-related writing were significant. Our more parsimonious final model is shown in Fig. 1.

We found that the final latent prior use variables had a direct positive effect of 0.03 ($p < 0.001$) on the mean achievement score for teacher-reported and 0.10 ($p < 0.001$) for student-reported prior use (see Fig. 2, for the Stata SEM results and Table 3 for the correlation matrix of scaled score, mean score, prior use components, and demographic controls). Goodness of fit statistics included an RMSEA of 0.04 and a CFI of 0.96, which indicate acceptable model fit and an improvement over the initial model (see discussion in Schreiber et al., 2006). We found that these latent prior use variables had a direct positive effect on the scaled achievement score of 0.07 ($p < 0.001$) for teacher-reported and 0.09 ($p < 0.001$) for student-reported prior use. Goodness of fit statistics were similar, with an RMSEA of 0.04 and a CFI of 0.95.

Using jackknife weighting, the coefficients for the mean score teacher-reported latent variable was 0.03 ($p < 0.05$) and student-reported school writing was 0.11 ($p < 0.001$) for a total of 0.14. The coefficients for the scaled score analysis were 0.07 ($p < 0.001$) for teacher-reported writing and 0.09 ($p < 0.001$) for student-reported writing, for a total of 0.16. See Table 2 for these results.

Using the latent variables of teacher and student-reported use of computers for writing for school and the dependent variable of mean individual scores, we ran a pooled booklet analysis on each of the 22 writing tasks controlling for demographics (without jackknife weighting). There was significant variability across the writing tasks, with the direct effect of teacher-reported school writing ranging from 0.05 ($p < 0.05$) to 0.10 ($p < 0.001$) and student-reported school writing ranging from 0.07 ($p < 0.01$) to 0.15 ($p < 0.001$). Please refer to Table 3 for full task-level results. Our SEM results show modest variation between analyses done at the booklet level and aggregated findings, with 17 out of 22 of the teacher-reported effect sizes and 15 out of 22 of the student-reported effect sizes of the booklet analyses falling within the confidence intervals of the aggregated findings (0.01–0.04 and 0.09 to 0.12, respectively).

4.2. OLS regression

Our regression analysis, with controls and interactions, using scaled and mean writing scores, also found positive effect sizes for student-reported prior use: 0.07 ($p < 0.001$) for the scaled achievement variable and 0.08 ($p < 0.001$) for the mean achievement scores. Teacher-reported writing effects were not statistically significant once controls were added into the regressions. These numbers are somewhat lower than those found in the SEM analysis, which would be expected. SEM allows stronger predictive power because measurement error is assumed to be a random error. This results in estimates of the path coefficients that are usually larger than if we had assumed no error in predictors, as with traditional regression models. Booklet-level results show an effect of student-reported writing ranging from 0.05 ($p < 0.05$) to 0.20 ($p < 0.01$) and similar interaction figures as the aggregated data, discussed below (see Table 4). Table 4 sets out the results of the aggregate regression analysis, which is consistent with our finding of a small positive relationship between prior student-reported use of computers for school writing and achievement on the 2011 NAEP writing assessment.

Table 2
Final structural equation model factor loadings, using jackknife weighting.

Latent variable	Coefficient	Standard error	z	P>|z|	95% confidence interval	
Mean Score Analysis						
Student-reported school writing	0.11	0.01	7.72	0.00	0.08	0.14
Teacher-reported writing	0.03	0.01	2.41	0.02	0.01	0.05
Aggregate	0.14	0.03				
Scaled Score Analysis						
Student-reported school writing	0.09	0.02	5.01	0.00	0.06	0.13
Teacher-reported writing	0.07	0.01	5.09	0.00	0.04	0.09
Aggregate	0.16	0.03				

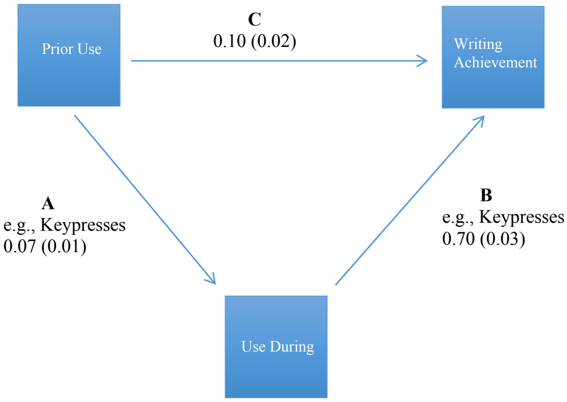

Fig. 2. Final partial mediation model. Computer use during the test (e.g., the number of keypresses) is impacted by prior computer use, computer use during the test predicts writing achievement, and prior computer use has an independent direct effect on writing achievement.

The only statistically significant interactions noted were slight, with teacher-reported use and parent's college education showing 0.02 ($p < 0.01$) effect on the scaled variable and 0.03 ($p < 0.01$) effect on the mean variable; student-reported prior use and free/reduced lunch status having -0.03 ($p < 0.001$) effect on both variables, student-reported prior use and current ELL status having mixed results depending on the achievement variables; and student-reported use and students with disabilities showing -0.04 ($p < 0.001$) effect on the scaled variable and -0.03 ($p < 0.01$) effect on the mean variable. Thus, the slightly positive benefits of prior computer use may be somewhat amplified if the student's parents have gone to college, and slightly reduced for students who are eligible for free/reduced lunch, currently designated ELLs, or are students with a disability. These interactions are small and not consistent across all of the achievement measures, thus warranting further research.

4.3. *Factor analysis*

We conducted a factor analysis to confirm that our latent variables were properly constructed using the complete group of 28 teacher and student survey questions. Principal factor analysis (unrotated) revealed that four factors had eigenvalues over 1 (Kaiser, 1960), and a significant decrease in eigenvalues for the subsequent factors (a drop from 1.51 to 0.67, suggesting a reasonable break point, see Tables 5 and 6). The underlying questions tended to relate to four topics: teacher-reported classroom use of computers for writing tasks ("Class Use," Factor 1, with eigenvalue of 3.93); student reported use of school-related computer-based writing ("Student Use," Factor 2, with eigenvalue of 3.15); teacher professional development in technology and instruction ("Teacher Development," Factor 3, with eigenvalue of 2.21) and teacher use of technology for writing instruction ("Teacher Use, " Factor 4, with eigenvalue of 1.501).

Checking the rotated model (oblique), we found similar eigenvalues, ranging from 3.17 to 1.55 for the first 10 factors (see Table 5 for the details of the oblique rotation eigenvalues and factor loadings). Factor loadings were sizable for the student-reported school writing questions (loadings of 0.70–0.85 for Factor 1), teacher-reported student use (0.60–0.79 for Factor 2), and student-reported home use (0.62–0.69 for Factor 3). Thus, the rotated factor analysis supported our three constructs: (a) student-reported school use of computers, (b) student-reported home use of computers, and (c) teacher-reported student use of computers. While teacher development was also supported as a construct in the unrotated analysis, it was not in the rotated analysis.

4.4. *Quasi-longitudinal check*

As a final check of our theory that computer-based writing benefits from practice writing *on computers* in a manner different from writing *on paper*, we used OLS regression to examine the relationship between prior computer use and achievement in a prior (paper-based) NAEP writing assessment.

The correlation between prior use and writing scores in 2007 is very small, 0.07, compared with the 2011 results that show a correlation of 0.19. Correlations between prior use and the other variables in both years are quite small, generally in the range of 0.01–0.08, with the higher correlations negative and relating to socioeconomic indicators in most cases (with the exception of similar correlations in this higher end of this negative range for current English language learners. Please refer to Table 6 for the correlation results.

By creating dichotomous variables for each level of prior use, we were able to test if the mean achievement was significantly different across each level of prior use. T-tests showed that in 2007 the difference in the means for students across these levels of prior use were statistically significant, but small in magnitude. The mean for students "never or hardly ever"

Table 3
SEM results, writing tasks (1–22) with individual mean scores as dependent variable, controls and no weighting.

Writing task	Sample size	Teacher-reported effect	Student-reported effect	RMSEA	CFI
1	2220	0.02 (0.02)	0.14 (0.02)***	0.04	0.95
2	2270	0.00 (0.00)	0.07 (0.02)**	NA	NA
3	2210	0.02 (0.04)	0.09 (0.03)**	0.04	0.96
4	2260	0.03 (0.02)	0.10 (0.02)***	0.04	0.96
5	2230	0.02 (0.02)	0.15 (0.02)***	0.04	0.96
6	2260	0.06 (0.02)**	0.08 (0.02)***	0.04	0.95
7	2250	0.04 (0.02)	0.03 (0.02)	0.04	0.96
8	2230	0.03 (0.02)	0.02 (0.02)	0.04	0.96
9	2250	0.10 (0.02)***	0.08 (0.02) **	0.04	0.96
10	2230	0.00 (0.00)	0.10 (0.02)***	NA	NA
11	2200	0.05 (0.02)*	0.13 (0.02)***	0.04	0.95
12	2240	0.06 (0.02)**	0.08 (0.02)***	0.04	0.96
13	2210	0.00 (0.00)	0.23 (0.02)***	NA	NA
14	2250	0.05 (0.02)*	0.11 (0.02)***	0.04	0.96
15	2230	0.04 (0.02)	0.07 (0.02)**	0.04	0.95
16	2270	0.03 (0.02)	0.17 (0.02)***	0.04	0.96
17	2240	0.06 (0.02)**	0.10 (0.02)***	0.04	0.96
18	2230	0.02 (0.02)	0.11 (0.02)***	0.04	0.96
19	2220	0.05 (0.02)*	0.11 (0.02)***	0.04	0.95
20	2260	0.03 (0.02)	0.09 (0.02)***	0.04	0.95
21	2220	0.05 (0.02)	0.02 (0.02)	0.04	0.96
22	2260	0.00 (0.00)	0.21 (0.02)***	NA	NA

Note: Sample sizes rounded to the nearest 10. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

using the computer for writing was −0.08, "sometimes" was −0.08, and "almost always" was 0.12. On the other hand, in 2011 students who never used the computer for school writing had a mean of −0.48, or 0.52 below students who sometimes or always used computers ($p < 0.001$). Thus, prior use of computers is associated with over one-half of a standard deviation increase in the computer-based writing scores of 2011. One sample t-tests confirm that the different means at each level in 2007 and 2011 are statistically significant. The prior use of computers has a small association with writing achievement in 2007, and in fact at low levels is showing a slightly negative relationship with scores. This is consistent with prior research suggesting that occasional use of computers, as compared to regular use, can be distracting for students, especially when preparing for paper-based testing (see discussion in Warschauer, 2006).

4.4.1. Regressions

We also ran an OLS regression of prior use on the scaled writing score, controlling for demographic measures. Table 7 shows the regression of scaled writing scores for the 2011 and 2007 assessments on the single variable relating to prior use. In 2011, each additional level of prior use (e.g., never to sometimes) was associated with a 0.17 standard deviation increase in the scaled writing score with all controls included. In 2007, an additional level of prior use is only associated with a 0.04 standard deviation increase. Although these numbers are all statistically significant, the 2007 results show that prior computer use had little practical effect on writing achievement scores in this assessment. The increase in scaled score in 2011 is four times that of 2007. The difference between the two effect sizes further confirms that prior use had a much larger effect on writing scores in 2011 when the assessment was computer based than it did in 2007 when the assessment was given on paper.

5. Discussion and conclusions

5.1. Prior use

Students' use of computers for school-related writing increased writing achievement on the NAEP computer-based writing assessment. Use of computers for other purposes, such as writing emails or writing blogs on the Internet, had no significant impact on achievement. Personal and home use of computers for unrelated matters did not increase writing achievement as measured by the NAEP assessment; whether or not the teacher used computers to model computer-based writing had only minimal effect on writing achievement as measured by the NAEP assessment. The benefit of prior computer use is slightly decreased for students who receive free/reduced price lunch or who have a disability. Other interactions were not consistent across the scaled and mean writing achievement variables and warrant further investigation.

Thus, as commonsense as it sounds, if schools want to increase the ability of students to write on computers, they need to provide more opportunities for the students to write for school on computers. Practice matters. Students will become more proficient at writing skills across modalities to use in their future college and career settings, allowing them to more fully participate in the community at large.

Table 4
Aggregate regression analysis, with controls and interactions, using scaled and mean writing scores.

	Scaled	Mean
Teacher-reported writing	0.01	0.01
	(0.01)	(0.01)
Student-reported writing	0.07***	0.08***
	(0.01)	(0.01)
Female	0.41***	0.45***
	(0.01)	(0.02)
Black	−0.42***	−0.46***
	(0.02)	(0.03)
Hispanic	−0.15***	−0.15***
	(0.02)	(0.03)
Asian	0.11***	0.16***
	(0.03)	(0.05)
Other	−0.10	−0.12
	(0.09)	(0.16)
Free/reduced lunch	−0.30***	−0.33***
	(0.01)	(0.02)
Parent college graduate	0.18***	0.18***
	(0.01)	(0.02)
Former ELL	−0.09*	−0.12*
	(0.03)	(0.06)
Current ELL	−0.72***	−0.57***
	(0.03)	(0.05)
Student w disability	−0.76***	−0.74***
	(0.02)	(0.04)
Teacher/female	0.00	−0.01
	(0.00)	(0.01)
Teacher/Black	0.00	−0.01
	(0.01)	(0.01)
Teacher/Asian	0.01	0.03
	(0.01)	(0.02)
Teacher/Hispanic	0.01	−0.01
	(0.01)	(0.01)
Teacher/free lunch	0.01*	0.01
	(0.01)	(0.01)
Teacher/parent college	0.02**	0.03**
	(0.01)	(0.01)
Teacher/current ELL	0.00	−0.01
	(0.01)	(0.02)
Teacher/former ELL	0.00	0.00
	(0.01)	(0.02)
Teacher/student w disability	−0.01	0.00
	(0.01)	(0.01)
Student/female	−0.01	0.00
	(0.00)	(0.01)
Student/Black	−0.02**	0.00
	(0.01)	(0.01)
Student/Asian	0.01	−0.01
	(0.01)	(0.02)
Student/Hispanic	−0.01	−0.01
	(0.01)	(0.01)
Student/free lunch	−0.03***	−0.03***
	(0.01)	(0.01)
Student/parent college	0.01*	−0.01
	(0.00)	(0.01)
Student/current ELL	−0.04***	0.08***
	(0.01)	(0.02)
Student/former ELL	0.01	0.00
	(0.01)	(0.02)
Student/student w disability	−0.04***	−0.03**
	(0.01)	(0.01)
Constant	0.04**	2.69***
	(0.01)	(0.02)
N	18,340	18,330
R-sq	0.33	0.15

Note: N rounded to the nearest 10. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 5
Eigenvalues and factor loadings for principal factor analysis (rotated, oblique) of computer use in the NAEP teacher and student surveys.

Factor	Variance	Proportion
Factor 1	3.17	0.03
Factor 2	2.81	0.27
Factor 3	2.35	0.23
Factor 4	2.26	0.22
Factor 5	2.14	0.20
Factor 6	2.00	0.19
Factor 7	1.97	0.19
Factor 8	1.67	0.16
Factor 9	1.45	0.14
Factor 10	1.15	0.11

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Uniqueness
Student, school, internet	**0.70**	−0.01	0.05	0.00	0.01	−0.02	0.02	−0.04	0.03	0.00	0.47
Student, school 1st draft	**0.71**	−0.03	0.04	−0.01	−0.01	−0.02	0.02	0.01	0.05	−0.03	0.46
Student, school, changes	**0.85**	−0.01	−0.05	0.00	0.01	0.01	−0.01	0.01	−0.02	0.01	0.33
Student, school, complete	**0.78**	0.05	−0.04	−0.01	−0.01	0.01	−0.02	0.02	−0.05	0.02	0.40
Student, school, write	0.36	0.05	0.39	−0.02	−0.01	0.01	−0.01	0.02	0.04	−0.03	0.55
Student, not school, write	0.01	−0.02	**0.62**	0.00	−0.01	−0.00	0.01	−0.02	0.03	−0.00	0.60
Student, not school, email	−0.02	0.00	**0.69**	0.01	−0.01	−0.00	−0.00	0.02	−0.01	0.00	0.54
Student, not school, internet	−0.01	0.00	**0.63**	0.00	0.02	0.01	−0.01	−0.01	−0.03	0.01	0.62
Teacher instruction, desktop	0.00	0.12	−0.00	0.05	0.39	0.00	0.01	0.02	−0.01	0.03	0.77
Teacher instruction, laptop	0.01	0.02	0.01	0.32	−0.10	0.09	−0.10	0.20	0.19	0.12	0.71
Teacher instruction, tablet	−0.00	−0.04	0.01	0.15	0.109	0.00	0.04	0.07	0.00	0.36	0.77
Teacher instruction, projector	−0.00	0.02	0.00	**0.69**	0.06	−0.04	0.03	−0.01	−0.01	−0.02	0.47
Teacher instruction, CD/DVD	0.01	0.00	−0.02	0.24	0.32	0.15	−0.08	0.05	0.17	0.05	0.60
Teacher instruction, digital device	0.03	−0.02	−0.02	0.09	0.20	0.03	−0.00	0.09	0.15	0.30	0.64
Teacher instruction, TV	−0.01	−0.02	0.00	−0.06	0.26	0.01	0.05	0.12	0.18	0.30	0.63
Teacher instruction, digital content	−0.02	0.02	0.02	0.28	0.25	−0.05	0.05	0.10	0.22	0.02	0.56
Teacher instruction, computer available	0.00	0.11	0.00	−0.02	0.020	−0.01	0.04	0.31	0.01	−0.01	0.87
Teacher instruction, internet available	−0.01	0.09	0.01	−0.11	−0.05	0.01	−0.02	−0.10	0.07	0.23	0.89
Teacher, students draft/revise on computer	0.02	**0.74**	−0.01	−0.00	0.02	0.01	0.02	0.13	−0.03	−0.02	0.36
Teacher, students complete writing on computer	0.02	**0.79**	−0.01	0.03	0.02	0.02	−0.03	0.05	−0.07	0.01	0.33
Teacher, students use word processing	0.00	**0.78**	−0.00	0.05	−0.02	0.02	−0.02	0.00	−0.03	−0.00	0.39
Teacher, students use internet for writing	0.00	**0.60**	0.00	0.03	0.07	−0.03	0.05	−0.03	0.14	0.02	0.55
Teacher, use computer for instruction	−0.02	0.12	0.01	0.65	0.07	−0.06	0.04	−0.01	0.01	−0.01	0.48
Professional development, basic computer	−0.02	0.00	0.01	−0.07	0.04	0.69	0.01	−0.01	0.01	−0.00	0.51
Professional development, software	0.01	−0.01	−0.01	0.03	−0.04	0.40	0.36	0.01	0.00	−0.01	0.51
Professional development, internet	0.01	0.01	0.00	0.02	−0.02	0.68	0.06	−0.01	−0.02	−0.02	0.49
Professional development, other technology	0.00	0.01	0.00	0.02	0.00	0.04	**0.68**	−0.00	0.01	0.01	0.51
Professional dev., integrating technology	0.01	−0.01	−0.00	−0.01	−0.02	0.04	**0.69**	0.03	−0.00	0.02	0.48

Factor rotation matrix

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Factor 1	0.56	0.71	0.39	0.53	0.57	0.16	0.19	0.58	0.46	0.34
Factor 2	0.77	−0.21	0.67	−0.42	−0.44	−0.14	−0.13	−0.30	−0.30	−0.30
Factor 3	0.02	−0.17	0.05	−0.15	−0.07	0.90	0.88	−0.03	−0.04	0.01
Factor 4	0.00	−0.63	0.30	0.51	0.33	−0.05	−0.05	0.10	0.35	0.18
Factor 5	−0.01	−0.01	0.01	−0.18	0.29	−0.34	0.33	−0.01	−0.11	0.12
Factor 6	−0.29	0.13	0.54	−0.17	−0.09	0.05	−0.07	−0.10	0.14	0.21
Factor 7	0.07	−0.02	−0.12	−0.42	0.09	0.16	−0.24	−0.17	0.52	0.73
Factor 8	0.01	0.05	−0.01	0.12	−0.27	−0.03	0.04	−0.47	0.05	0.34
Factor 9	0.00	−0.02	0.01	0.03	0.10	0.02	−0.02	0.12	−0.52	0.19
Factor 10	−0.00	−0.03	−0.01	0.01	−0.43	−0.04	0.03	0.55	0.13	0.14

Correlation matrix of the oblimin(0) rotated common factors

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Factor 1	1.00									
Factor 2	0.19	1.00								
Factor 3	0.57	0.01	1.00							
Factor 4	−0.01	0.16	0.03	1.00						
Factor 5	0.01	0.29	−0.03	0.56	1.00					
Factor 6	0.00	0.03	0.00	−0.03	0.02	1.00				
Factor 7	0.02	0.04	0.01	0.06	0.14	0.69	1.00			
Factor 8	0.11	0.37	0.02	0.53	0.39	0.07	0.15	1.00		
Factor 9	0.03	0.19	0.09	0.33	0.39	0.17	−0.08	0.27	1.00	
Factor 10	−0.05	0.21	0.01	0.09	0.34	0.17	−0.04	0.09	0.64	1.00

Note. Rotated factors are correlated. Method: principal factors. Rotation: Oblique Oblimin (Kaiser on). Observations: 22,150. Retained factors = 10. Number of parameters: 235. LR test: independent vs. saturated: chi2(378) = 2.0e + 05 Prob > chi2 = 0.0000.
Bolded loadings are those above 0.40.

5.2. Implications

The affordances of computers as writing tools and the amount that these affordances lead to both improved quality and increased quantity of writing has been highly variable in research to date (Applebee & Langer, 2011; Bangert-Drowns, 1993; Cochran-Smith, 1991; Collins et al., 2014; Goldberg, Russell, & Cook, 2003; Graham & Perin, 2007; Morphy & Graham, 2012; Russell & Haney, 1997; Russell & Plati, 2002; Schwartz & Bridwell, 1984; Zheng, Warschauer, Lin, & Chang, 2016). Computer-based writing is often implemented in schools with contextual changes that support improved writing more generally—both on and off computers. For example, the affordances of digital tools encourage increased collaboration, authentic writing audiences, meaningful tasks, mentoring, and motivation (Warschauer, 2011).

The fact that school-related digital writing had a greater impact on the students' achievement on the NAEP assessment than did personal, casual computer use is consistent with the large research base showing the need for technology to be implemented in instructionally sound ways (see discussion in Warschauer, 2011). Simply putting digital tools in students' hands will not improve their learning (OECD, 2015). Rather, the tools must be integrated in a way that supports and extends the curriculum in meaningful ways (Warschauer, 2011).

Our analyses confirm the preliminary findings of a recent working paper commissioned by NCES of three NAEP digitally based assessments in math (2011), writing (2011), and technology and engineering literacy (2013); eighth graders' self-reported familiarity with and prior use of digital technology positively impacted their scores on the assessments (Zhang et al., 2016). Their factor analysis of the student surveys revealed two distinct types of prior computer use, use for school-related writing and use for more general writing (e.g., emails, personal writing; Zhang et al., 2016). Regressed on students'

Table 6
Correlation of scaled writing score, frequency of computer use for writing (single question variable), and demographic variables, 2011.

	Scaled score	Prior use	Female	Hispanic	Black	Asian	Free/Red lunch	Parent HS grad	ELL current	ELL former	Student w Disab
Scaled Score	1.00										
Prior Use	0.19* (0.00)	1.00									
Female	0.24* (0.00)	0.07* (0.00)	1.00								
Hispanic	−0.16* (0.00)	−0.07* (0.00)	0.00 (0.66)	1.00							
Black	−0.20* (0.00)	0.01 (0.18)	0.00 (0.61)	−0.27* (0.00)	1.00						
Asian	0.08* (0.00)	0.07 (0.00)	0.00 (0.68)	−0.13* (0.00)	−0.11* (0.00)	1.00					
Free/Red Lunch	−0.34* (0.00)	−0.11* (0.00)	0.01 (0.29)	0.28* (0.00)	0.25* (0.00)	−0.05* (0.00)	1.00				
Parent HS Grad	0.14* (0.00)	0.09* (0.00)	−0.05* (0.00)	−0.28* (0.00)	0.06* (0.00)	0.03* (0.00)	−0.25* (0.00)	1.00			
ELL Current	−0.23* (0.00)	−0.04* (0.00)	−0.02* (0.00)	0.26* (0.00)	−0.07* (0.00)	0.07* (0.00)	0.17* (0.00)	−0.17* (0.00)	1.00		
ELL Former	−0.05* (0.00)	−0.01 (0.13)	0.00 (1.00)	0.27* (0.00)	−0.008* (0.00)	0.04* (0.00)	0.13* (0.00)	−0.12* (0.00)	−0.05* (0.00)	1.00	
Student w Disab	−0.28* (0.00)	−0.04* (0.00)	−0.08* (0.00)	−0.01 (0.51)	0.03* (0.00)	−0.05* (0.00)	0.07* (0.00)	−0.02* (0.00)	0.03* (0.00)	−0.01 (0.23)	1.00

Correlation of scaled writing score, frequency of computer use for writing (single question variable), and demographic variables, 2007

	Scaled score	Prior use	Female	Hispanic	Black	Asian	Free/R lunch	Parent HS Gr	ELL current	ELL former	Student w Disab
Scaled Score	1.00										
Prior Use	0.07* (0.00)	1.00									
Female	0.25* (0.00)	0.02* (0.00)	1.00								
Hispanic	−0.13* (0.00)	−0.05* (0.00)	0.02* (0.00)	1.00							
Black	−0.16* (0.00)	−0.03* (0.00)	0.012* (0.00)	−0.18 (0.00)	1.00						
Asian	0.05* (0.00)	0.03* (0.000)	−0.01* (0.04)	−0.08* (0.00)	−0.10* (0.00)	1.00					
Free/Red Lunch	−0.27* (0.00)	−0.09* (0.00)	0.02* (0.00)	0.25* (0.00)	0.28* (0.00)	−0.02* (0.00)	1.00				
Parent HS Grad	0.13* (0.00)	0.07* (0.00)	−0.04* (0.00)	−0.27* (0.00)	0.03* (0.00)	0.02* (0.00)	−0.23* (0.00)	1.00			
ELL Current	−0.23* (0.00)	−0.04* (0.00)	−0.02* (0.00)	0.26* (0.00)	−0.07* (0.00)	0.07* (0.00)	0.17* (0.00)	−0.17* (0.00)	1.00		
ELL Former	−0.02* (0.00)	−0.001* (0.00)	0.01* (0.01)	0.20* (0.00)	−0.04* (0.00)	0.05* (0.00)	0.08* (0.00)	−0.10* (0.00)	−0.02* (0.00)	1.00	
Student w Disab	−0.21* (0.00)	−0.01* (0.00)	−0.09* (0.00)	−0.01 (0.00)	0.03* (0.00)	−0.03* (0.00)	0.08* (0.00)	−0.03* (0.00)	0.02* (0.00)	−0.02 (0.00)	1.00

Note. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.005$.

Table 7
Regression of scaled writing scores, frequency of computer use for writing (single question variable), demographic variables, 2011

	Scaled writing score					(Parent college)		
Prior Use	0.28***	0.25***	0.23***	0.19***	0.18***	0.16***	0.18***	0.17***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female		0.43***	0.44***	0.45***	0.45***	0.41***	0.44***	0.41***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black			−0.64***	−0.42***	−0.44***	−0.45***	−0.44***	−0.45***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Hispanic			−0.53***	−0.32***	−0.26***	−0.17***	−0.17***	−0.19***
			(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Asian			0.06*	0.10***	0.12***	0.13***	0.20***	0.15***
			(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Other			omitted	omitted	omitted	omitted	omitted	omitted
Free/Red Lunch				−0.47***	−0.43***	−0.32***	−0.40***	−0.37***
				(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Parent HS Grad					0.20***		0.15***	0.14***
					(0.02)		(0.02)	(0.02)
Parent College Grad						0.20***		
						(0.01)		
Former ELL						−0.10**	−0.09**	−0.10*
						(0.03)	(0.03)	(0.03)
Current ELL						−0.71***	−0.73***	−0.70***
						(0.3)	(0.03)	(0.03)
Student w Disability						−0.75***		−0.75***
						(0.02)		(0.02)
Constant	−0.66***	−0.82***	−0.52***	−0.32***	−0.48***	−0.30***	−0.42***	−0.32***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
N	23,850	23,850	23,220	21,950	19,760	19,760	19,760	19,760
R-sq	0.04	0.09	0.19	0.23	0.23	0.31	0.25	0.30

Regression of scaled writing scores, frequency of computer use for writing (single question variable), demographic variables, 2007

	Scaled writing score					(Parent college)		
Prior Use	0.10***	0.09***	0.07***	0.05***	0.04***	0.03***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female		0.48***	0.49***	0.49***	0.49***	0.45***	0.49***	0.45***
		(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Black			−0.49***	−0.31***	−0.33***	−0.33***	−0.33***	−0.34***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Hispanic			−0.49***	−0.30***	−0.24***	−0.17***	−0.15***	−0.17***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asian			0.01	0.07***	0.09***	0.10***	0.14***	0.11***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Other			−0.50***	−0.33***	−0.33***	−0.26***	−0.29***	−0.28***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Free/Red Lunch				−0.40***	−0.36***	−0.27***	−0.35***	−0.31***
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Parent HS Grad					0.27***		0.24***	0.22***
					(0.01)		(0.01)	(0.01)
Parent College						0.21***		
						(0.01)		
Former ELL						−0.05*	0.01	−0.03
						(0.02)	(0.02)	(0.02)
Current ELL						−0.55***	−0.55***	−0.53***
						(0.01)	(0.01)	(0.01)
Student w Disability						−0.75***		−0.76***
						(0.01)		(0.01)
Constant	−0.22***	−0.45***	−0.24***	−0.10***	−0.31***	−0.10***	−0.29***	−0.19***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	138,020	138,020	136,630	132,010	117,210		117,200	117,200
R-sq	0.01	0.07	0.13	0.16	0.16		0.17	0.23

Note. N rounded to the nearest 10. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.005$.

writing achievement (using plausible values, rather than scaled or mean scores), the researchers found that both types of activities led to positive increases in writing scores, but that school-related computer writing was much more predictive of higher writing scores, controlling for gender, ethnicity, free/reduced lunch status, and urbanicity (but not controlling for parent education, ELL, or special education status; Zhang et al., 2016).

Given the prevalence of computer-based writing in the world beyond the school gates, we believe that school writing should at least in part consist of digital writing. Schools are situated to provide the necessary access, instruction, and support to enable students to become proficient writers on computers. Because disparities in access to technology and the Internet remain significant, both at home and at school, improving school use of computers for writing can help reduce the digital divide. We note, however, that certain sub-groups of students show reduced improvement from the prior use of computers. These differences are worth additional research to ensure that efforts to improve students' digital writing do not increase the divide.

Despite the lack of statistical findings with respect to teacher's use of technology to provide writing instruction, teachers should still be encouraged to incorporate technology into their lessons. Tools that make writing visible, by the teacher, the student, and peers, still provide useful instruction. They are also increasingly the way writing is done in professional and academic settings, with collaboration becoming increasingly important. Similarly, teachers' need for quality professional development in integrating technology into quality curriculum remains despite the lack of a direct statistically significant effect in our analysis.

Two areas need further investigation: (a) what are the reasons for the differences in results depending on the use of scaled or mean writing scores and (b) what are the reasons for the variability in our booklet compared to aggregated analyses. Our analysis found a larger booklet-to-aggregate results variation than found in the Horkay et al. (2006) analysis, which warrants further analysis. We suspect that there may be a link to the writing genre involved in the task or the order of presentation of the task (first versus second question answered), though preliminary results are not clear.

Finally, we think that it would be helpful to look at more specifics of how prior use is improving writing achievement, both quantitatively and qualitatively. The NAEP assessment collected keystroke data, which may further illuminate how and for whom prior use impacts achievement (e.g., do students with more prior use tend to delete more). We expect to find patterns of both productive and unproductive keystrokes. In addition, case studies observing adolescents actually in the process of writing on computers should give us a better understanding of the writing process and how it progresses and, perhaps, changes as students become more proficient writing on computers. Computers will have both positive and negative affordances for writing, and different students will navigate those affordances differently.

5.3. Limitations

5.3.1. NAEP assessment

The nature of our writing achievement variable, based on the NAEP assessment, is inherently limited. The NAEP assessment measures only two 30-min writing sessions. The time limit means that the writing samples are rough drafts and not polished final versions. By design, the NAEP assessment is not reflective of students' abilities to edit and refine their work. The time limit advantages students who are used to writing for similar lengths of time. The time limit may disadvantage students with language production disabilities or English-language learners who could use additional time, but additional time could frustrate other students and create fatigue (see Applebee, 2007). In addition, the functionality of the NAEP interface used in the 2011 test could be improved, as seen during the usability studies conducted in 2012 with fourth grade students (NCES, 2014a, b). As these functionality improvements are made in future years, we might find that the interface is simpler and easier to use for students with less prior exposure to computers, which could in turn reduce the correlation of prior use with writing achievement.

5.3.2. Variables

This analysis is limited to modeling the effect of student-reported prior computer use on writing achievement. Future analyses will also consider use of the computer during the assessment itself and the relationship between use during the assessment and achievement, as well as the interaction between prior computer use and use of the computer during testing on achievement. Simple reported frequency of use does not speak to the quality of instruction in computer-based writing, nor is student-reported frequency as accurate as real-time measures of computer use might be. Finally, we note that our research intentionally does not address total prior time spent writing or quality of writing instruction received, which we expect would be more directly related to writing achievement than computer usage information.

In conclusion, systematic analysis of the 2011 NAEP writing test scores demonstrates that frequency of prior computer-based writing in school is moderately correlated with computer-based writing achievement. This adds weight to the argument that increased integration of technology in K-12 education is required if we are to prepare students fairly for a future of computer-based writing.

Acknowledgements

This work has been supported by funding from the Spencer Foundation, Grant No. 201500153.

References

Abedi, J. (2002). A latent-variable modeling approach to assessing reliability and validity of a creativity instrument. *Creativity Research Journal, 14*(2), 267–276.

Allen, N. L., & Donoghue, J. R. (1996). Applying the Mantel-Haenzel Procedure to complex samples of items. *Journal of Educational Measurement, 33*(2), 231–251.

Applebee, A. N. (2007). Issues in large-scale writing assessment perspectives from the national assessment of educational progress. *Journal of Writing Assessment, 3*(2), 81–98. Retrieved from http://www.journalofwriting assessment.org/archives/3-2.2pdf.
Applebee, A. N., & Langer, J. A. (2009). What is happening in the teaching of writing? *English Journal, 98*(5), 18–28.
Applebee, A. N., & Langer, J. A. (2011). A snapshot of writing instruction in middle schools and high schools. *The English Journal, 100*(6), 14–27.
Bangert-Drowns, R. (1993). The word-processor as an instructional tool: A meta-analysis of word processing in writing instruction. *Review of Educational Research, 63*(1), 69–93.
Bazerman, C. (2011). Standpoints: The disciplined interdisciplinarity of writing studies. *Research in the Teaching of English, 46*(1), 8–21.
Beaton, A., Rogers, A., Gonzalez, E., Hanly, M., Kolstad, A., Rust, K., et al. (2011). *The NAEP primer*. Washington, DC: National Center for Educational Statistics.
Beaton, A. E., & Zwick, R. (1992). Overview of the national assessment of educational progress. *Journal of Educational and Behavioral Statistics, 17*(2), 95–109.
Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research*. NY: Guilford Press.
Bowen, W. G., Kurzweil, M. A., & Tobin, E. M. (2005). *Equity and excellence in American higher education*. Charlottesville, VA: University of Virginia Press.
Cochran-Smith, M. (1991). Word processing and writing in elementary classrooms: A critical review of related literature. *Review of Educational Research, 61*(1), 107–155.
College Board. (2015). *SAT college-bound seniors total group profile report*. Retrieved from https://secure-media.collegeboard.org/digitalServices/pdf/sat/total-group-2015.pdf_September 6, 2015.
Collins, P., Hwang, J. K., Zheng, B., & Warschauer, M. (2014). Writing with laptops: A quasi-experimental study. *Writing & Pedagogy, 5*(2), 203–230.
De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*(4), 687–698.
Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Tatum, J. (2008). *Cognitive models of writing: Writing proficiency as a complex integrated skill*. Princeton, NJ: Educational Testing Service.
DeBono, T., Husseini, A., Cairo, C., Ghelani, K., Tannoch, R., & Toplak, M. E. (2012). Written expression performance in adolescents with attention-deficit/hyperactivity disorder (ADHD). *Reading & Writing, 25*, 1403–1426.
Englert, C. S., Mariage, T. V., & Dunsmore, K. (2006). Tenets of sociocultural theory in writing instruction research. In C. A. MacArthur, S. Graham, & J. Fitzgereald (Eds.), *Handbook of writing research* (pp. 208–221). NY: Guilford.
Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992–2002. *The Journal of Technology, Learning, and Assessment, 2*(1).
Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.
Haertel, E. (1984). An application of latent class models to assessment data. *Applied Psychological Measurement, 8*(3), 333–346.
Horkay, N., Bennett, R. E., Allen, N., 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).
Jackman, M. R., & Jackman, R. W. (1983). *Class awareness in the United States*. Berkeley, CA: University of California Press.
Johnson, E. G. (1992). The design of the national assessment of educational progress. *Journal of Educational Measurement, 29*(2), 95–110.
Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.
Langer, J. A. (1991). Literacy and schooling: A sociocognitive perspective. *Literacy for a Diverse Society: Perspectives, Practices, and Policies*, 9–27.
Lyon, G. R., & Weiser, B. (2013). The state of the science in learning disabilities: Research impact on the field from 2001 to 2011. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 118–151). NY: Guilford.
Messick, S., Beaton, A., & Lord, F. (1983). *National assessment of educational progress reconsidered: A new design for a New era*. Princeton, NJ: Educational Testing Service.
Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Read Write, 25*, 641–678.
National Assessment Governing Board (NAGB). (2010). *Writing framework for the 2011 national assessment of educational progress*. Washington, D.C.: U.S. Department of Education.
National Center for Education Statistics. (2012). *The Nation's report Card: Writing 2011*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
National Center for Education Statistics. (2014a). *Lessons learned from the 2012 Grade 4 writing computer-based assessment study*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/nationsreportcard/writing/lessons/default.aspx.
National Center for Education Statistics. (2014b). *NAEP 2012 pilot writing computer-based assessment study—grade 4: Technical memorandum*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
National Center for Education Statistics (NCES. (2009). *The Nation's report Card: An overview of procedures for the NAEP assessment*. Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.
Organisation for Economic Co-operation and Development (OECD). (2015). *Students, computers and Learning: Making the connection*. Retrieved from http://dx.doi.org/10.1787/9789264239555-en.
Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco, CA: Jossey-Bass.
Rogers, A., Stoeckel, J., & Sikali, E. (2013). *National assessment of educational progress (NAEP) 2011 writing assessment restricted-use data files data companion*. Washington, D.C.: National Center for Education Statistics.
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), 1–18.
Russell, M., & Plati, T. (2002). Does it matter with what I write? Comparing performance on paper, computer and portable writing devices. *Current Issues in Education, 5*.
Saenz, V. B., Hurtado, S., Barrera, D., Wolf, D., & Yeung, F. (2007). *First in my family: A profile of first-generation college students at four-year institutions since 1971*. Los Angeles, CA: Higher Education Research Institute.
Sandene, B., Horkay, N., Bennett, R. E., Allen, N., Braswell, J., Kaplan, B., et al. (2005). *Online assessment in Mathematics and Writing: Reports from the NAEP technology-based assessment project*. Washington, DC: National Center for Education Statistics.
Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs* (report from the Governing Board of the American Educational Research Association Grants Program). Washington, DC: American Educational Research Association.
Schreiber, J., Nora, A., Stage, F., Barlow, E., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Education Research, 99*(6), 323–338.
Schwartz, H. J., & Bridwell, L. S. (1984). A selected bibliography on computers in composition. *College Composition and Communication, 35*(1), 71–77.
Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417–453.
Snibbe, A. C., & Markus, H. R. (2005). You can't always get what you want: Educational attainment, agency, and choice. *Journal of Personality and Social Psychology, 88*, 703–720.
Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), *The concept of activity in soviet psychology* (pp. 144–188). Armonk, NY: M. E. Sharpe.
Warschauer, M. (2006). *Laptops and literacy: Learning in the wireless classroom*. NY: Teachers College Press.
Warschauer, M. (2011). *Learning in the cloud: How (and why) to transform schools with digital media*. NY: Teachers College Press.
Welch, W. W., Anderson, R. E., & Harris, L. J. (1982). The effects of schooling on mathematics achievement. *American Educational Research Journal, 19*(1), 145–153.
Wenglinsky, H. (2005). *Using technology wisely: The keys to success in schools*. NY: Teachers College Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.
Wertsch, J. V. (1998). *Mind as action*. Oxford: Oxford University Press.
Zhang, T., Xie, Q., Park, B. J., Kim, J., Yee, Broer, M., et al. (2016). *Computer familiarity and its relationship to performance in three NAEP digital-based assessments* (Retrieved from Washington, DC).
Zheng, B., & Warschauer, M. (2015). Participation, interaction, and academic achievement in an online discussion environment. *Computers & Education, 84*, 78–89.
Zheng, B. B., Warschauer, M., Lin, C.-H., & Chang, C. (2016). Learning in one-to-one laptop environments: A meta-analysis and research synthesis. *Review of Educational Research, XX*(X), 1–33. http://dx.doi.org/10.3102/0034654316628645.
Zwick, R. (1987). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement, 24*(4), 293–308.