Educational Evaluation and Policy Analysis
March 2016, Vol. 38, No. 1, pp. 65–87
DOI: 10.3102/0162373715592706

Improving Access to, Quality, and the Effectiveness of Digital Tutoring in K–12 Education

Patricia Burch
University of Southern California
Annalee Good
University of Wisconsin-Madison
Carolyn Heinrich
University of Texas at Austin

There is considerable variation in how providers of digital education describe what they do, their services, how students access services, and what is delivered, complicating efforts to accurately assess its impact. We examine program characteristics of digital tutoring providers using rich, longitudinal observational and interview data and then analyze student attendance patterns and effects of digital tutoring on low-income students' reading and mathematics achievement. We find significant associations between formats, curriculum drivers, tutor locations, and other characteristics of digital providers and their effectiveness in increasing student achievement, as well as differential access by student characteristics, that warrant further investigation as digital providers' roles in K–12 instruction continue to expand.

Keywords: *digital instruction, tutoring, student achievement*

Digital instruction—using a digital platform (such as computer, netbook, or handheld device) as an integral and consistent part of an instructional delivery strategy—is rapidly becoming a commonplace component of K–12 classroom and supplemental instruction. Estimates place the current value of the U.S. market for K–12 education software and digital content anywhere in the range of US$8 billion (Molnar & Cavanaugh, 2014) to US$21 billion per year (Burch & Good, 2014). In the last decade, private-sector investment in K–12 education technology companies has nearly tripled, from US$146 million to US$420 million (Burch & Good, 2014). As of 2011, 63% of districts with enrollments higher than 10,000 students contracted with an outside organization to provide online courses (Queen, Lewis, & Coopersmith, 2011). Advances in technology have allowed digital tools to compete with features of face-to-face instruction with the promise of low-cost, broad access (Richards & Struminger, 2013).

In this research, we focus on digital providers' role in out-of-school time (OST) tutoring programs, which has continued to expand, even as waivers from No Child Left Behind (NCLB) have released many districts from the requirement to offer federally funded supplemental education services. In a mixed-method, longitudinal study of OST tutoring conducted in five urban sites over 4 school years, we observed online tutoring companies reaching a student "market share" as high as 88% in one district; in another district, we observed a single digital provider delivering tutoring to more than 10,000 students. NCLB mandated unfettered parental choice in tutoring providers and accordingly gave providers the flexibility to try varied formats for tutoring. However, the implementation and effects of the wide range of approaches and formats that are emerging in digital tutoring are especially difficult for school districts to monitor and assess.

Moreover, there is considerable variation in how digital tutoring providers describe what they do, the actual services they offer, how students access these services, what is delivered, and the degree of alignment to state standards and district needs, which complicates efforts to accurately assess the effects of digital tutoring on students' academic achievement. Drawing on our 4-year mixed-methods study of federally funded OST tutoring programs, we examine key program characteristics of digital providers, as described in provider applications for state approval, recorded in district administrative data and enacted and observed in rich, longitudinal observational data. Specifically, we ask what are the key characteristics of different program models in digital tutoring (curriculum, instructional driver, the role of the tutor, use of data, etc.), as reflected in program descriptions in state applications, in district administrative data, and in observational data of instructional settings. We identify critical variables that define the format and content of digital tutoring, as well as access points for students enrolled in digital tutoring (e.g., location of tutor and curriculum). We then conduct exploratory analyses of student attendance patterns and the relationship of different digital provider characteristics, tutoring forms, and access points to the educational outcomes (reading and math achievement) of students from low-income families.

Our findings raise concerns about which students have access to the types or forms of digital tutoring that the results suggest may be relatively more effective. We find that English-language learners and students with disabilities were significantly less likely to receive OST tutoring in formats that value-added models suggested may be more effective in increasing student math achievement. Based on these findings, we consider priority directions for research that aims to improve digital tutoring models, and the policy tools available to state and local educational agencies to ensure greater transparency and continuous improvement of the quality of digital tutoring and its accessibility (Miron & Urschel, 2012).

Prior Research

There is a growing demand for more and more rigorous evidence to understand whether and how "digital" and "tutoring" practices in K–12 systems are linked to student achievement outcomes (Cavanaugh, Barbour, & Clark, 2009). The few studies examining the effects of different kinds of digital instruction on student outcomes show mixed results (Bingham, in press; Burch & Good, 2014), and they seldom focus on the K–12 student population in the United States (London, Pastor, & Rosner, 2008; Price, Richardson, & Jelfs, 2007; Slattery, 2003). In addition, while recent studies have started to build a knowledge base on the characteristics of quality digital instruction (more generally), to date, equity issues have received less attention in the literature. This is concerning given the implicit suggestion in some studies that online instruction has distinct advantages for students who are economically and academically disadvantaged (Rose & Blomeyer, 2007). In this section, we review what is known about the types of digital instruction associated with quality instruction and student achievement gains. Next, we motivate the importance of greater treatment of equity issues in research on instructional technology and, in doing so, set the context for subsequent discussion of digital instruction in federally funded OST programs in this study.

The existing research on digital education and student learning is limited, particularly in the context of increasing calls for expanding education technology in public schools. (Bingham, in press; Means, Bakia, & Murphy, 2014). A handful of studies have found positive effects linked to specific online formats. For example, Bakia, Shear, Toyama, and Lasseter (2012) found that blended (both online and in-person) instruction can lead to positive effects on student achievement, especially when it is collaborative and promotes self-reflection in students. Arroyo, Tai, Muldner, Woolf, and Park (2013) found digital mathematics instruction to be particularly beneficial to female students' mathematics knowledge and problem-solving ability. Other studies emphasize the importance of live interaction between teachers and students for improving educational outcomes (Zhao, Lei, Yan, Tan, & Lai, 2005), as well as real-time data feedback for teachers and consistent access to the technology for all students, regardless of need (Brush & Hew, 2006).

Alternatively, some researchers have found no effect or negative effects of blended learning models on student achievement (Cole, Kemple, & Segeritz, 2012; Margolin, Kieldon, Williams, & Schmidt, 2011). A study of School of One's Math-Only blended learning program based in New York City examined achievement gains of School of One students, comparing them with the achievement gains of School of One students prior to the blended learning intervention. The School of One study controlled for prior achievement, student demographics, and city and statewide factors. On average, researchers found that the School of One blended model did not improve sixth graders' math achievement. The lack of effect was explained in terms of a "gap dip," where students were filling in gaps in their knowledge instead of working on grade-level skills. Similarly, a study of the Enhancing Education Through Technology Program in Vermont drew on survey data, interviews, and site visits to evaluate the program's implementation, technology integration, sustainability, and perceived effects on student outcomes (Margolin et al., 2011). Findings from this study identified classroom organization as a major challenge for teachers implementing the program. The program was organized to enable students to work at their own pace; however, some students had difficulty working independently and teachers lacked capacity to effectively organize the classrooms in ways that supported independent learning. Reviewing the literature on virtual schools, Barbour and Reeves (2009) found a mix of both benefits to student learning (e.g., higher levels of student choice and motivation) and challenges (e.g., retention and a lack of access associated with the "digital divide").

A critical but relatively overlooked issue underlying the extant research on digital learning concerns the extent to which digital instruction addresses long-standing inequities and achievement gaps. This is a pressing concern, as districts and states are increasingly requiring some form of online instruction as a condition of graduation (Burch & Good, 2014). For example, among the students eligible for OST tutoring in our study, anywhere from two thirds to 100% are free-lunch eligible, 90% to 98% are students of color, and up to 36% are English language learners. Historically, these students are some of the most vulnerable in terms of achievement gaps, and there is growing evidence that students in poverty still face considerable barriers to accessing products and services offered under the banner of digital education (Goslee & Conte, 1998; Zickuhr & Smith, 2012). That is, as the use of technology in public education expands, access to this technology is lower for students attending schools with a higher percentage of families living in poverty (Burch & Good, 2014; Snyder & Dillow, 2013).

Clearly, more rigorous research on the effects of digital tutoring in K–12 settings is needed. At the same time, if this research is to inform the rapidly expanding policy and program agendas that are encouraging online instruction, there also needs to be more specific attention to understanding the attributes of digital tutoring that work for students with varying levels and types of instructional needs, as well as the capacity required of large, urban school districts for managing the use of educational technology.

Our mixed-methods study of the implementation and effects of digital providers in federally funded OST tutoring is intended to contribute to the knowledge base on digital programs and practices in OST settings. In light of the limited evidence on online tutoring, we leverage the larger research base identifying factors that contribute to high-quality OST tutoring in traditional bricks-and-mortar settings to inform our work. These studies suggest that a high-quality OST curriculum is content rich, differentiated to student needs, and connected to students' school day (Beckett et al., 2009; Stevens, 2012; Vandell, Reisner, & Pierce, 2007). Effective instruction is organized into small grouping patterns (ideally 3:1 or less), and instructional time is consistent and sustained. Instructional strategies are varied, active, focused, sequenced, and explicit (Beckett et al., 2009; Elbaum, Vaughn, Hughes, & Moody, 2000; Farkas & Durham, 2006; Lauer et al., 2006; Little, Wimer, & Weiss, 2008, Lou et al., 1996; Vandell et al., 2007). And beyond elements specific to curriculum and instruction, quality OST programs not only hire and retain tutors with both content and pedagogical knowledge but also provide instructional staff with continuous support and feedback (Little et al., 2008; Vandell et al., 2007). Research also suggests the importance of actively supporting positive relationships among tutors and students (Durlak & Weissberg, 2007; Durlak, Weissberg, & Pachen, 2010; Vandell et al., 2007), as well as between programs and the surrounding community (Little et al., 2008).

We argue that there is a need for more research on how these best practices in OST tutoring hold or diverge in digital OST settings. The research and findings we present below link information on digital instruction formats and other program attributes and their implementation in OST settings with data on student achievement to explore the effects of digital OST on student achievement, including for subgroups targeted by NCLB.

Research Samples, Methods, and Data

This investigation builds on a longitudinal, mixed-method study of OST tutoring, including an in-depth, qualitative examination of instructional practice in different program models and settings and a rigorous, quasi-experimental analysis of OST tutoring program effects. The study sample includes students eligible for OST tutoring under NCLB[1] in five urban school districts—Chicago, Milwaukee, and Minneapolis Public Schools; Dallas Independent School District (Dallas ISD); and Los Angeles Unified School District—that ranged in size from approximately 80,000 to 650,000 students over the study period, 2009 to 2013 (see Table 1). Student demographics in these districts are generally representative of the larger national population that is eligible for OST tutoring, that is, high concentrations of students from low-income, urban settings, including subgroups with higher levels of academic need/disadvantage (e.g., students with limited English proficiency and disabilities). Our study data also include information on approximately 180 providers of OST tutoring in these districts, about a quarter of which are digital providers.

Although we draw on both quantitative and qualitative data collected in this study for the 2009–2010 to 2012–2013 school years, we describe the research we present here as primarily an exploratory effort to dig deeper into the "black box" of digital instruction in the OST context. A key aim of this work is to examine the characteristics of digital OST instruction (and its management and implementation by providers and districts) and to develop a conceptual framework that links them to improvements in student learning and achievement. Our qualitative investigation draws on data collected within and across the study districts described above to identify key program attributes and practices of digital OST tutoring, while our quantitative analyses of digital instruction focus on a single school district (Dallas ISD) for which detailed coding of digital provider characteristics was undertaken (and linked to information on students served by these providers). Dallas ISD provided scans of the applications that OST tutoring providers submitted to the state of Texas to obtain approval for offering OST tutoring in the district, as well as administrative data that included information about the instructional settings, tutor location, tutoring format (e.g., individual, small group, etc.), tutoring subject, student–teacher ratios, and digital access points. These data were combined and analyzed to construct the detailed measures of digital OST program features that we use in our empirical analysis of tutoring effects on student achievement. Table 2 presents descriptive statistics on the students eligible, registered for and, attending OST tutoring in Dallas ISD in 2011–2012 (the year for which we have detailed data on digital OST providers), as well as these same statistics for the students among these who were matched with digital OST providers.

We have also conducted in-depth, qualitative observations and collected other data on 32 OST tutoring providers in our multisite study, including seven digital tutoring providers. The sample of seven digital providers is illustrative of the key subcategories of digital program formats that we further discuss below, including: synchronous (live), asynchronous (not live), entirely digital, and blended (digital and in-person), as well as both national and locally based providers. Four of these seven digital providers serve a market share of 14% or higher in at least one of our study districts. For the purposes of analysis, a "digital" provider is one that uses a digital

TABLE 1

Characteristics of Students Eligible for OST Tutoring in Study Districts

Characteristics	Chicago public schools				Dallas independent school district				Los Angeles unified school district		Milwaukee public schools				Minneapolis public schools			
Year	2008–2009	2009–2010	2010–2011	2011–2012	2008–2009	2009–2010	2010–2011	2011–2012	2010–2011	2011–2012	2008–2009	2009–2010	2010–2011	2011–2012	2008–2009	2009–2010	2010–2011	2011–2012
Number of students	88,353	87,542	101,930	245,616	35,612	30,774	35,026	39,091	344,323	364,837	11,992	26,798	16,439	20,905	10,963	15,769	16,444	15,906
Asian (%)	1	2	2	2	1	1	1	1	2	3	5	4	4	4	11	9	9	9
Black (%)	53	49	42	43	34	31	30	31	8	8	68	69	68	68	48	47	46	45
Hispanic (%)	44	47	53	51	62	64	65	62	85	82	17	20	20	20	28	29	28	26
White (%)	2	2	2	3	3	4	4	3	3	5	8	5	8	6	6	8	9	12
Other race (%)	0	0	1	1	0	0	0	1	2	3	3	3	0	0	6	7	7	8
% female	49	49	49	50	48	48	48	48	49	49	48	47	46	46	51	50	50	49
% ELL	12	12	16	18	21	19	16	20	31	28	6	10	12	10	34	36	33	36
% free lunch	100	100	100	99	67	79	74	60	100	100	83	87	88	90	99	100	100	98
% with disabilities	14	13	12	13	12	12	11	11	10	10	21	22	22	24	17	18	18	18
Attended SES last year (%)	26	42	8	13	16	15	37	28	10	11	11	6	14	8	13	7	16	16
% absent last year	6	4	5	5	7	9	7	6	5	4	16	15	16	13	8	8	7	4
Retained this year (%)	4	2	2	2	0	7	8	8	4	4	13	11	12	12	2	6	2	5

Note. OST = out-of-school; ELL = English language learners; SES = Supplemental Educational Services.

TABLE 2

Characteristics of Students Eligible, Registered for and Attending OST Tutoring in Dallas Independent School District and Matched With Digital Providers

	Dallas independent school district, 2011–2012 school year					
	District eligible sample			Students matched with digital providers		
	Eligible	Registered	Attended	Eligible	Registered	Attended
Number of students and characteristics	39,091	10,862	7,941	11,111	7,610	5,651
Asian (%)	1	0	1	1	1	1
Black (%)	31	33	33	30	29	28
Hispanic (%)	62	64	64	65	68	68
White (%)	3	2	2	3	2	2
Other race (%)	1	1	1	1	0	1
% female	48	49	49	49	49	49
% ELL	20	23	24	23	25	26
% free lunch	60	84	84	74	84	84
% with disabilities	11	12	13	12	11	12
Attended SES last year (%)	28	37	39	40	40	41
% absent last year	6	5	5	6	4	4
Retained this year (%)	8	5	4	6	5	5
Middle school (%)	30	30	31	28	30	31
High school (%)	67	69	68	70	69	68

Note. OST = out-of-school time; ELL = English language learners; SES = Supplemental Educational Services.

platform (software or live tutor via technological platform such as computer, netbook, or handheld device) as an intentional, integral, and consistent part of its instructional strategy in delivering tutoring to eligible students in at least one of the five districts in our study. Students served by these providers consistently used digital instructional tools for at least half of their tutoring experience.

In undertaking the qualitative work, we used a standardized observation instrument in both nondigital and digital tutoring settings (Burch & Good, 2013). Because digital and nondigital settings can differ in a number of ways, this instrument includes indicators that specifically accommodate digital settings without a live tutor (e.g., instructional software that adapts to students' instructional needs), as well as measures that describe how technology is used to improve instruction (e.g., to use higher order thinking skills) and to address issues of access (e.g., reliability and accessibility to all students). Over 4 years, we observed 185 full tutoring sessions (46 across the seven digital providers in our study sample). Other elements of the qualitative data collection in the larger study include 79 personal interviews with provider administrators about the structure of instructional programs, choice of curricula and assessments, challenges in implementation, and choices in staffing; 109 personal interviews with tutoring staff about instructional formats, curriculum, adaptations for students' learning differences, staff professional background, and training; 47 personal interviews with district and state administrators involved in program implementation; focus groups with 221 parents/guardians of students who were eligible to receive OST tutoring, and document analysis of formal curriculum materials from providers; diagnostic, formative, or final assessments used; and informal "in-use" curriculum collected during instructional sessions and policy documents on federal, state, or district policies concerning program implementation.

In the quantitative analysis, our empirical measures of the key treatment variables—hours of OST program participation, types of digital and nondigital tutoring, and hours/types of combinations—are constructed using district administrative data and the qualitative data collected and coded to describe digital tutoring features and formats. The administrative data from the school districts allow for the construction of dosage measures of tutoring with specific providers. Specifically, the tutoring providers were required to invoice school districts for each hour of tutoring provided to the students, and thus, tutoring "dosages" are measured in invoiced hours of tutoring (per student). Other data made available by the districts included the rate per hour charged by the providers, the total of invoices paid out, and, in some cases, the balance of unspent funds (from dollars allocated per student for tutoring). The data from the digital provider applications and other components of the qualitative research investigation aided in developing empirical measures of variables such as tutoring formats/types and forms of digital tutoring.

We also obtained student-level demographic, attendance, and test score data from the school districts. These data include controls for gender, race/ethnicity, free and reduced-price lunch eligibility, English proficiency, students with disabilities, grade retention, prior year achievement test scores, number of absences from the prior school year, grade year, school attended, and prior OST tutoring program attendance (see the descriptive statistics in Table 1). These are standard, student-level control variables (in an education production function model). In addition, student outcomes—specifically, student test scores on state standardized tests—are measured as effect sizes, that is, the level of student achievement relative to the district average score on state standardized tests. These achievement measures are derived from student test scores on the Texas Assessment of Knowledge and Skills (TAKS) state standardized test, which was used in determining adequate yearly progress (AYP) under NCLB.

Qualitative Analysis

Data analysis in the qualitative component of this study occurred both concurrent to and after the data collection process, using a constant comparative method to explore and explain provider instructional practices. Analytic codes were developed from patterns in initial data collection and in response to the research questions, and then reapplied to interview, observation, and archival data to establish findings. Coding trees and data were inputted into a qualitative coding system where researchers collaborate on common project tasks through remote access to a common server. The base and examples of associated subcodes applied to qualitative fieldwork include "enrollment" (e.g., process, strategies, challenges), "instructional core" (e.g., amount of instructional time, differentiation, curriculum structure and/or source, varied instruction, classroom-level interaction, tutor capacity), "alignment" (e.g., individualized learning plans, instructional practice, challenges), and "students with special needs" (e.g., areas of confusion, curriculum, instruction, format, challenges). The research team then developed additional subcodes specific to digital provider analyses, which included "technology–instructional format," "technology–curriculum," "technology–assessment," "technology–access," "technology–administrative uses," and "technology–free" to capture relevant data on technology outside of those subcodes. For purposes of analysis, all audiotapes of focus groups and interviews were transcribed verbatim and then transformed into integrated text for analysis.

In addition to coding the text recorded in observations, ratings of indicators were analyzed by categorizing indicators into clusters, organized by areas of OST tutoring best practice (e.g., varied, active, rigorous, targeted, differentiated, high levels of student engagement). This clustering of qualitative indicators allowed us to see which best practices are predominant in observations and which were rare or missing. For example, in assessing whether a session was "active," we would focus on indicators such as whether students had to participate in structured discussions, demonstrate understanding of concepts, or help determine the direction of an instructional task. Levels of differentiation were examined in terms of indicators such as accommodations made for students with disabilities or English language learners, or whether a software program or tutor adapted the instructional pace or content based on student needs. In assessing the rigor of a tutoring session, we would include indicators that focused on the extent to which instructional tasks required by curriculum software and/or the tutor demanded the application of students' higher order thinking skills, or if students were asked "why," "how," or "what if" questions as part of the session. Although the observation instrument ratings used a numeric rating system, the process was fully qualitative in terms of clustering the indicators under each best practice area.

Quantitative Analysis

In the larger study of OST tutoring on which we build this work, we have used multiple strategies for quasi-experimental estimation of OST tutoring impacts, including value-added modeling, fixed-effects models (student fixed effects and student plus school fixed effects), and propensity score matching methods. We have found a high degree of consistency in the estimates produced by these models (Heinrich & Nisar, 2013), and therefore have primarily used a value-added modeling approach that controls for school fixed effects.

In estimating the relative effectiveness of different features/formats of digital tutoring, our sample for estimation consists of all students receiving digital tutoring, and we adjust for selection into different types of digital providers. For each estimation, we make the assumption that after adjusting for all available measured characteristics and prior test scores, program participation (i.e., receipt of a particular type of digital tutoring) is independent of the student outcomes that would occur in the absence of participation (in a particular type of digital tutoring). We also recognize, however, that there could be factors (for which we do not have measures and do not control for in our models) that could explain both participation in specific types of digital programs and student outcomes, leading to possible bias in our estimates of digital provider effectiveness. For example, we do not have complete information on the extent to which tutoring providers may have influenced student enrollment in their programs with promises of access to digital devices (or particular types of devices), and whether this type of information may have encouraged selective enrollment among students with differing levels of access to or experience with digital tools in school and/or home settings, which might in turn have affected the extent to which students made academic progress through use of digital tools. In the absence of concrete information on how selection may have worked in this regard—for example, how students with less experience versus more experience with digital tools differentially chose among the digital provider options, or how important of a factor was this in their decisions—it is difficult to speculate on the direction of any potential omitted variable bias.

The particular value-added model (with school fixed effects) that we use allows us to control for other classroom and school interventions which are fixed over time. For example, if there is a reading intervention at a school and those students also receive tutoring in that program, failing to control for the intervention (school fixed effect, π_s) would bias the results. We estimate

$$A_{jst} = \alpha DigCharac_{jt} + \beta X_{jt-1} + A_{jst-1} + \pi_s + \mu_{gt} + \varepsilon_{jst}, \quad (1)$$

where A_{jst} is the achievement of student j attending school s in year t; DigCharac$_{jt}$ is an indicator function if the student j attended tutoring with a digital provider with a given characteristic in year *t;* X_{jt-1} are student characteristics which include student demographics, percent absent in prior year, retained in prior year, and attended tutoring in the prior year; A_{jst-1} is the prior year test score; π_s is school fixed effect; μ_{gt} are grade by year fixed effects; and ε_{jst} is the random error term. Identification in this specification comes from average student achievement after controlling for student characteristics and school and grade year effects. In these models, we include one or more indicators of digital program characteristics, as all students in these analyses will have received tutoring from a digital provider. The outcome measure is the level of student (math or reading) achievement, adjusting (on the right-hand side) for the possibility that students with similar characteristics might enter OST tutoring with different underlying achievement trajectories (as reflected in their prior test scores).

Because our value-added modeling strategy includes school-level fixed effects, we are utilizing the within-school variation in attributes of the OST program offerings to identify any effects of digital program characteristics on student achievement. Our data analysis confirms that there is substantial within-school variation in the distribution of the OST program characteristics (described in greater detail below), specifically, variation in the presence of (and combinations of) characteristics that include the location of the tutor, instruction drivers, curriculum location, and tutor synchronicity. The exhaustive descriptive analysis (available from the authors) showed that only one characteristic—having a "tutor-structured curriculum driver"—was not present among the providers delivering OST tutoring to students in 3 of the 26 schools.

Our quantitative analysis is tightly linked with the qualitative research in defining measures, specifying the empirical models and analyzing the factors that influence the outcomes of digital OST tutoring. For example, as detailed below, interviews and observations from the qualitative fieldwork revealed important differences within digital tutoring formats, critical information that was then applied in refining our measures and interpretation of empirical results. We also optimized our sample through this integrated mixed-methods approach by using quantitative data to identify the parameters (e.g., student market share, cross-site enrollment, etc.) that guided the selection of tutoring providers observed in the field research. We think that this tightly integrated, mixed-methods approach strengthens the validity of the inferences from this exploratory work.

Research Findings

Indicators of Instructional Quality

In addressing the quality of digital OST tutoring and constructing our measures of quality, we drew upon two sources of observation data: average ratings on select indictors and narrative description of tutoring sessions, both captured on the standardized observation instrument. In addition to observation data, we also drew from interviews in identifying key elements of the digital OST settings.

TABLE 3
Comparison of Observation Ratings for Select Indictors of Instructional Quality (2009–2013)

Indicator	Digital	Nondigital
Ask students why, how, or what if questions	.24	.52
Challenge students to push themselves intellectually	.30	.50
Students push themselves intellectually	.29	.51

Table 3 offers rating averages of three primary indicators of instructional rigor. Although indicators on the instrument are rated from 0 to 2 (with a 2 meaning that it was observed consistently throughout the observation point with most students), the averages here are recorded from 0 to 1, where a "1" would indicate that an indicator received a rating score of "2" in every observable instance. Comparing the average rating of digital and nondigital tutoring sessions across 4 years (2009–2013), the digital tutoring sessions were rated low overall as well as in comparison with nondigital settings.[2] More specifically, these average indicators suggest that digital OST sessions lacked important elements of high-quality instruction, such as intellectual rigor and the application of higher order thinking skills. Average ratings across at least 50 observation points indicate that digital sessions were even less effective at encouraging these elements than the already low ratings for nondigital sessions.

In addition, we added three pilot indicators in the last year of data collection (2012–2013), specific to the digital setting. Table 4 presents data from 25 observations across five of the digital providers in four of our study districts. Again, the averages below are recorded from 0 to 1, where a "1" would indicate that an indicator received a rating score of "2" in every observable instance.

The juxtaposition of additional narrative elements from observations of instructional settings with these ratings offers a further perspective on the quality of digital OST tutoring. For example, as shown in Table 4, technology was generally reliable and accessible to students participating in the settings we observed. When we did see difficulties with accessing the instructional material, it

TABLE 4

Observation Ratings for Digital-Specific Indicators (2012–2013)

Pilot indicator	Average rating
Technology used is reliable and accessible to all students	.78
Instructional software adapts to students' needs	.30
Use technology to employ higher order thinking skills	.16

related to either problems initially logging in or with audio equipment associated with synchronous (live) tutoring. To mitigate technical problems, one provider in our qualitative sample held training sessions with parents and students before the commencement of services. This involved a 2-hour session to introduce the curriculum, what instruction was going to look like, and how to use the laptops for instruction.

Two of the providers used a program where students moved independently through preloaded or Internet-accessed curriculum software without a live tutor present. This presented a challenge to students who might get stuck on a problem. However, where providers (four in our sample) combined face-to-face tutoring with online software, tutors had the capacity to differentiate the instruction and reword some of the existing problems. Alternatively, for those providers using a live tutor, we observed few instances where the instructor changed a full problem. The tutors sometimes asked students to draw representations of the problem on a digital whiteboard during math instruction, but only to help explain the problem or as a way for the instructor to see a student's work. In three of the four synchronous providers, instructors rarely provided any follow-up questions or any differentiation aimed at simplifying a question or increasing the level of difficulty. In three of the providers where tutors worked with multiple students at once via the online platform, students had to wait for the instructor to give them the next problem. Students who finished early had to wait about 2 to 3 minutes to move ahead, while the tutor was helping other students in the virtual classroom.

As the data in Table 4 also indicate, there was little evidence of the use of technology to use higher order thinking skills. Often, the questions presented to students were simply "digitized worksheets" that did not require students to actually use technology to apply, evaluate, or create concepts. In general, our preliminary analysis of tutoring practices across different digital providers suggests that digital tutoring, not surprisingly, does not always add value to instructional quality, even when the technology is working well and is accessible on-site.[3]

As part of our ongoing, mixed-methods efforts to better understand the quality of digital OST tutoring, we have identified three elements of digital tutoring that offer a critical vantage point on the levers for improving the quality of instruction in the digital setting. These include (a) the nature of curriculum and what drives it, (b) what drives the instruction and the role of the tutor, and (c) the nature and role of assessments and data in digital tutoring programs.

Digital Curriculum

Due to regulations under NCLB, the general content focus of many digital providers in the OST context is either language arts or math. However, providers, whether under the law or operating in states with waivers from NCLB, are given considerable discretion in how they enact the curriculum, contributing to considerable variation in terms of curricular format, curricular access, and curricular software. Curricular formats range from highly structured and completely dependent upon software to "homegrown" curriculum that is more fluid and dependent on the discretion of a live tutor. For example, one provider uses software that is essentially an online whiteboard through which the tutor and student interact by writing with the track pad/mouse, typing, and speaking through headsets. The tutor can upload curriculum materials and prompts as needed. In terms of source, curriculum used by digital providers comes from a variety of sources (purchased/leased from an outside source to curriculum developed in house and used only by tutors, and some combination of above). A number of providers develop their own, proprietary curriculum used only by their tutors. We find that digital curriculum used in tutoring is often delivered outside of the traditional classroom and school context, so that teachers and principals are unable to do a "walk through" to observe curriculum and instruction. For that and other reasons, it is much harder to "see" and analyze particular types of curricula and, in particular, the enacted curriculum.

In addition, students access the curriculum in a variety of ways. In our qualitative sample of seven providers, we have seen one provider lend students used desktop computers, another provides a handheld device, two provide netbooks, and the remaining three providers send students laptop computers. Each of these providers had either software preloaded onto the hardware or dedicated websites through which students would access the program. All but two of these providers used Internet-based programs.

In a digital tutoring setting, software is a key element of the instructional setting. Drawing on both our qualitative investigation of the digital tutoring setting and common terms used in the field of digital education (iNACOL, 2011), we identify three types of software used to facilitate instructional interactions between students, tutors, and curriculum in our qualitative study sample:

- *Synchronous instructional software* facilitates live instructional interaction between students and tutors through chat functions, audio capabilities, and/or a "whiteboard" function. This type of software houses the curricular content itself and in principle is intended to generate progress reports.
- *Synchronous course management system* (CMS) facilitates live instructional interaction between students and tutors, for example, through a "whiteboard" platform combined with an Internet-based voice call service (e.g., Skype). This type of software facilitates digital interaction between the student and the tutor, but the tutor generates or delivers "homegrown" curricular content.
- *Asynchronous instructional software* houses curricular content but does not support live interaction between students and tutors. This software may house assessments, generate progress reports, and use "artificial intelligence," in other words software developed to adapt the pace and direction of tasks based on student responses.

Instructional Driver and the Role of the Tutor

From our own and others' prior research, we know that the role of the tutor is key to instructional quality (Good, Burch, Stewart, Acosta, & Heinrich, 2014; Hock, Pulvers, Deshler, & Schumaker, 2001). The context of digital tutoring challenges traditional conceptions of a "tutor." Instead of falling into the models typical of in-person, nondigital tutoring contexts where the tutor is the primary guide or delivery system of the curriculum, our observations of tutoring sessions and interviews with provider staff indicate and illustrate a spectrum of enacted roles. For this analysis, we define "tutor" as the provider staff most directly responsible for the instruction of an individual student; in other words, the closest adult to the point of instructional delivery. We categorize digital OST tutors into the following:

- *No tutor:* Some digital tutoring platforms are structured where students have no interaction with a human during the tutoring session. Instead, students interact with instructional software, and may have the option of calling a helpline if they get stuck on a problem. Students also might interact with a provider staff member on occasions to upload progress reports or deal with technical needs (see below).
- *Technician:* Some tutoring platforms use personnel only for technical assistance, which could include a technical helpline or delivering/retrieving hardware from students' homes. We also observed sessions where students brought netbooks into a central location to have a provider personnel upload their progress in working through preloaded software.
- *Monitor/guide:* Tutor and "monitor" are beyond a technician, but not quite a full, interactive instructor. We characterize the "monitor" role as when tutors respond to students if they need help on a specific question related to academic content, call families to discuss progress and encourage students, or answer questions via email.
- *Instructor:* We identify a tutor as an instructor if the tutor interacts with a

student constantly throughout the session, and the curriculum could not progress without the tutor. The instructor category differs from that of the monitor/guide in that the tutor is an integral part of instructional platform and curriculum delivery.

Although these roles are distinct, in practice, tutors often occupy multiple categories, sometimes simultaneously. For example, we observed a synchronous tutoring session where the tutor was working through a math problem with a student when the audio connection with the student was lost. The tutor then had to use the chat function in the software program to explain how to reconnect the headset, so that they could resume instruction. In addition to tutors, there may be staff farther from the point of instructional delivery, but who interact with a student's instructional process. These include case managers, teacher leaders/monitors, curriculum managers, and so on: for example, counselors or case managers who contact parents and the school district if there are issues or questions about students' progress, or "prescription monitors" who periodically review student files, adjust the sequence or pace of the learning program, and continually train tutors. There are also provider staff involved in instructional delivery, but who do not interact with students or their files. These include, for example, curriculum teams that continue to develop and revise the curriculum, or quality assurance testers that test the curriculum once it is inserted into the software platform.

Use of Data and Assessments

Assessment and the data it generates are just as important of a consideration in digital tutoring as curriculum and instruction, and just as complex. The distinctions between curriculum, instruction, and assessment often blur, especially for those programs where the software drives the assessment, which drives aspects of the curriculum, which in turn drives instruction. Under NCLB, all OST tutoring providers, whether digital or nondigital, were required to provide pre- and posttest scores for every student in their program. Some districts offer or require the use of their own assessments as pre- and posttest (e.g., progress assessments given in the fall, winter, and spring). Other districts require providers to obtain and administer their own. For providers in our study, digital assessments were either developed by the provider in-house or purchased from another company, or the provider had access to district assessments for use as pre- and posttests. For those providers administering their own pre- and posttests, assessments were in a digital format, except in the case of one provider that conducted verbal assessments of kindergarten and first-grade students who might have problems navigating the digital platform.

All of the providers in our sample also used some type of formative assessment to measure progress and potentially revise the scope and sequence of a student's learning plan. These formative assessments were often short sets of problems designed to gauge whether students understood a concept. Some software would either not allow students to move forward unless they correctly responded to these problems, or a live tutor approved their progress and moved them to the next activity. What is very clear from our analysis at this point is that, as in nondigital tutoring, there is considerable variation in how digital OST tutoring providers describe what they do, the actual services they offer, how students access these services, and what is delivered.

Publicly Available Information on Instructional Setting

Our in-depth examination of the digital OST instructional setting described above offers important insights into some of these challenges of determining if and how digital tutoring affects student achievement. One of these critical insights is how different digital formats can be from one another in terms of how they are described by providers (the intended curriculum) in publicly available information, such as provider applications or parent brochures. For example, a provider may simply indicate that its program includes a particular type of software, but not specify whether it is used for pre- and postassessment or actual instruction. Based on analyses of provider applications to the state of Texas for offering services in Dallas ISD, we identify the following preliminary patterns in the types and quality of information provided to parents for choosing providers.

First, the information parents receive about vendor programs can be diluted and misleading. For example, a vendor might say that they provide services for students with disabilities but do not actually hire tutors with special education training. Vendor program descriptions often provide minimal information for parents on how they actually use technology as part of instruction. For instance, some provider applications made mention of the use of instructional websites, but a closer reading of the application indicated that only the tutors (and not the students) access these websites to gather curricular materials. Second, it is difficult to find a single, consistent source of program descriptions. On a number of occasions, the program description in providers' state applications differed from the description in district parent information. Third, some providers were described as having digital platforms 1 year but not the next. Fourth, there are many different types of digital platforms, which are often not specified in the application. Finally, there are providers that do not include digital tutoring as any part of their program description or marketing materials, but individual tutors may choose to include digital learning tools as part of the regular curriculum. One example is teachers in one district having the kids do part of their tutoring session on a classroom computer with the same instructional software program the district uses with all students in day school instruction.

Digging Deeper to Classify Dimensions of Digital Tutoring

Drawing on the analysis described above of the nature of curriculum, instruction, assessment, and information in enacted tutoring, we developed a categorization system for digital providers for use in rigorous estimation of OST tutoring program effects. The work of developing the new taxonomy for digital tutoring was done iteratively with the work of classifying digital tutoring programs based on the self-descriptions in their Texas provider applications. To specify the universe of our taxonomy, we defined "digital" tutoring services as those in which students directly interact with digital technology. For our classification purposes, we generally considered "digital technology" to be any multipurpose computer device at least as sophisticated as an iPod or other tool of equivalent functional capacity, which also includes tablets, netbooks, laptops, and desktop computers but does not include less versatile electronic tools such as digital calculators.

The complexity of the latter work—the application analysis and provider classification process—varied considerably among provider applications. Among the applications we analyzed, there were a number of reasons why a tutoring program's characteristics might have been hard to discern from the provider application. These classification challenges included inadequately framed or specified application questions, vague information in provider responses, insufficient details about program characteristics in provider responses, conflicting details about program characteristics in provider responses, and inconsistent degree of details on different modes of tutoring in provider responses (when providers offer multiple tutoring modes). In these cases, we not only had to iteratively refine our taxonomy while classifying providers' tutoring programs according to that taxonomy, but we also had to iteratively assess each tutoring program's actual characteristics for classification, while determining which application text excerpts were relevant for justifying those classifications and cross-checking them with available district administrative data on provider attributes.

Based on descriptive analysis of the applications of approved tutoring providers in Dallas ISD during the 2012–2013 school year, we developed the following categorization that both leverages and digs deeper into characteristics (instruction, curriculum, assessment) identified in observational work.

- Tutor location: Where does the student access the tutor?
  - Online or via the phone (remote access)
  - Face-to-face (in-person access)
- Tutor synchronicity: How immediate is the student's communication with the tutor?
  - Asynchronous (time-delayed)
  - Synchronous (live)
- Instruction driver: Who or what is guiding the student's learning?

  - Curriculum-based software (locally installed or delivered online)
  - Tutor actively working through curriculum-based software with the student
  - Tutor without curriculum-based software (often using a digital whiteboard if online)
- Curriculum location: Where does the student access the course content?
  - Via a digital device, over the Internet (includes mobile device software that needs ongoing Internet access to provide content)
  - Via a digital device, using locally installed software (includes mobile device software that does not need the Internet to provide content once installed)
  - Via nondigital resources (e.g., books, worksheets, chalk/whiteboard, etc.)[4]

In our analyses of Dallas ISD digital OST tutoring programs, we have used this structure and a set of categorizations to explore associations between digital provider and program attributes and student achievement. We summarize preliminary findings of the quantitative analysis below.

Preliminary Empirical Findings on Digital Provider Effects on Student Achievement

A primary objective of this empirical work was to explore the potential effects of different types of digital tutoring (and their delivery) that contribute to student achievement. The analysis of digital providers in Dallas ISD links the data extracted and coded (per the categories of digital tutoring described above) from the state applications of 35 digital providers (with the largest student market shares in the 2011–2012 school year) to district administrative data on digital providers and student-level data on 11,111 students served by these providers. We think it is important to emphasize again that these data are based in part of information self-reported from the digital providers, and thus, some caution is warranted in examining associations between digital provider attributes and student characteristics and achievement.

Tables 4 and 5 present basic descriptive information on the types of digital programs/providers and the proportions of students they enroll, as well as how hourly rates charged by the providers vary across formats/types. Using two-group mean comparison tests, cross-tabulations with chi-square tests, and logistic regression, we examined student selection into different types of digital providers, looking for associations between student characteristics and the provider characteristics as shown in Table 5. The strongest (statistically significant) associations we found (specifying $\alpha = .05$ and two-tailed tests) were for students with special needs and the instruction driver and tutor synchronicity attributes of providers. Specifically, both two-group mean comparison tests and cross-tabulations with a chi-square test showed that students with disabilities were more likely to be tutored with curriculum-based software ($p = .0256$) or a tutor with software combination ($p < .0001$), while English language learners were also more likely to receive OST tutoring through a combination of tutor and software-driven instruction ($p < .0001$). In addition, these descriptive tests showed that English language learners ($p < .0001$) and students with disabilities ($p = .0250$), as well as students of Hispanic origin ($p < .0001$), were significantly *less* likely to receive OST tutoring in synchronous formats.

The logistic regressions controlled for the same student characteristics as shown in Table 1 and predicted the probability of receiving tutoring from a digital provider with a given provider characteristic, as shown in Table 5. The results of these analyses confirmed the statistically significant associations found in the descriptive analyses and provided additional information on their magnitude. For example, the odds of a student with disabilities being tutored with curriculum-based software were 49% higher than for students without disabilities. And while we expect synchronous formats of tutoring to be more effective, the odds of Hispanic students receiving tutoring in this format were 34% lower than for non-Hispanics, and they were also 17% lower for English language learners and 20% lower for students with disabilities. These analyses also showed other interesting associations between student characteristics and digital program attributes, such as that students absent from school more often were significantly more likely to receive all of their tutoring online (with no

TABLE 5
Profile of Digital OST Tutoring Providers in Dallas Independent School District, 2011–2012

Digital provider characteristic	% of students (2011–2012)
Tutor location	
Entirely on Internet	6.36
All in-person	10.78
Face-to-face and online	82.84
Instruction driver	
Curriculum-based software	7.77
Tutor-structured	1.39
Tutor with curriculum-based software	7.29
Combination tutor with software-driven and tutor-driven	24.23
Software-driven and tutor-driven	52.88
Curriculum location	
Curriculum location only digital online	17.78
Curriculum location only nondigital	0.01
Digital-online and local-nondigital combination	60.09
Digital-online and digital-local combination	20.35
Tutor synchronicity	
Asynchronous	2.67
Synchronous	19.31
Combination of synchronous and asynchronous	78.00
Described as blended	2.50

Note. OST = out-of-school time.

face-to-face tutoring).[5] Although it is plausible that digital providers tailored some of their OST offerings to meet the special needs of particular subgroups of students, our qualitative field research showed that, more often than not, digital providers were not prepared to differentiate instruction to better serve students with special needs (i.e., lacking the information necessary to do so, such as student individual education plans, or the capacity, such as bilingual tutoring staff).

In our multisite study of OST tutoring, we found that digital providers, on average, charged significantly more per hour (about US$20 more per hour) than nondigital providers and delivered fewer hours of services to students than face-to-face tutoring providers. In Dallas ISD, the average hourly rate charged by digital providers (in the 2011–2012 school year) was US$31/hour higher than that of nondigital providers (US$89/hour vs. US$58/hour). Students attending with digital OST providers also received significantly fewer hours of tutoring (13 vs. 22 hours) on average (or 41% fewer hours). The information in Table 6 includes the hourly rates only for digital providers in our Dallas ISD subsample and shows how they varied by digital program characteristics. Interestingly, the results show that digital providers that combined digital online with face-to-face instruction were charging the highest rates per hour (in terms of tutor location). In addition, those that were combining some form of tutor-structured with software-driven curriculum were also charging the highest rates among the varying forms of instruction drivers. This same pattern follows for curriculum location and tutor synchronicity as well: Blending different attributes within a given digital provider is associated with higher hourly charges for services. This begs the question: Are these provider attributes that are linked with higher hourly rates also associated with student achievement in reading and/or math?

The results from the value-added models (with school fixed effects) that examine associations between digital provider characteristics and student achievement (in math and reading) are

TABLE 6
Provider Rates by Digital Characteristics (Reporting Statistically Significant Differences)

Tutor location	Rate (US$)
Entirely on Internet	55
All in-person	74
Face-to-face and online	88
Instruction driver	
Curriculum-based software	80
Tutor-structured	69
Tutor with curriculum-based software	62
Combination tutor with software-driven	88
Software-driven and tutor-driven	86
Curriculum location	
Curriculum location only digital online	70
Curriculum location only nondigital	
Digital-online and local-nondigital combination	92
Digital-online and digital-local combination	86
Tutor synchronicity	
Asynchronous	58
Synchronous	66
Combination of synchronous and asynchronous	90

shown beginning in Table 7, which focuses on tutor location (i.e., where the student accesses the tutor). In this estimation, we look at the relationship between tutor–student interactions that are entirely on the Internet or all in-person (face-to-face) versus the reference category of a blend of online and face-to-face and student achievement in math and reading. Table 7 also shows the coefficient estimates and robust standard errors for student-level controls, but for brevity, it does not present the coefficient estimates for the school fixed effects or the indicator variables that control for grade level.

This first set of results (see Table 7) suggests that students who receive OST tutoring from digital providers in which access to the tutor is all face-to-face potentially realize significantly larger benefits in terms of their math achievement (compared with those where the tutor location is a blend of online and face-to-face); the estimated effect is also more than 3 times the size of that for students receiving tutoring entirely on the Internet. In addition, the coefficient estimate for all in-person/face-to-face is substantively large relative to the average effect sizes of OST tutoring that have been reported in our larger study and related research, typically ranging from .05 to .10 standard deviations (Heinrich et al., 2014; Heinrich, Meyer, & Whitten, 2010; Heistad, 2007; Rickles & Barnhart, 2007; Springer, Pepper, & Ghosh-Dastidar, 2009; Zimmer, Gill, Razquin, Booker, & Lockwood, 2007; Zimmer, Hamilton, & Christina, 2010). In effect, the highest priced (in terms of provider hourly rates) tutor location (online/face-to-face blend) appears to be the least effective for tutoring in math. We see no statistically significant associations between tutor location and student reading achievement.

Table 8 presents the findings of value-added models that compare the effectiveness of alternative instruction driver forms (who or what is guiding the students' learning) in digital OST tutoring. The results again differ for math and reading. The least effective instruction driver for math OST tutoring is a combination of tutor-with software-driven and tutor-driven instruction (relative to tutor-driven and software-driven), which is also billed at the highest hourly rate on average. For reading, however, curriculum-based software instruction drivers are significantly less effective in increasing student achievement. Tutor-structured—where the tutor structures and drives the student's learning without curriculum-based software—is negatively associated with student math and reading achievement, although these and the other estimated effects of instruction drivers are not statistically significant.

With respect to curriculum location (where the student accesses the tutoring content), there is only one statistically significant association with student achievement—a negative association between math performance and curriculum that is a combination of digital-online and digital-locally accessed (see Table 9). This is in comparison with the reference category—a digital-online and local-nondigital combination—which is the most prevalent and also the most expensive location (in terms of provider hourly rates) where students access tutoring content. Finally, we also see (in Table 10) a statistically significant, positive

TABLE 7
Value-Added With School Fixed-Effects Models of Digital Provider Effects: Tutor Location

Digital provider and student characteristics	Math score (standardized)		Reading score (standardized)	
	Coefficient	*SE*[a]	Coefficient	*SE*
Tutor location				
Online/entirely on Internet	0.040	.075	−0.037	.052
All in-person/face-to-face[b]	**0.153**	.034	0.055	.043
Prior year standardized score	**0.335**	.067	**0.391**	.036
Attended OST tutoring last year	0.037	.033	0.039	.026
Asian	0.194	.359	0.062	.298
Hispanic	0.094	.059	0.093	.058
Other race	0.075	.193	0.100	.095
White	−0.052	.094	−0.042	.116
Free-lunch eligible	0.026	.031	**0.124**	.037
English language learners	**−0.160**	.065	−0.077	.062
Student with disability	0.020	.187	0.110	.197
Female	0.046	.038	**0.088**	.029
Percentage of days absent from regular school in prior year	**−1.703**	.770	**−2.812**	.683
Retained in grade	−0.139	.129	**−0.709**	.207
Constant	−0.034	.207	0.562	.283

Note. Additional controls (not reported): School fixed effects and grade year. Boldface indicates statistical significance at .05. OST = out-of-school time.
[a]Robust standard errors.
[b]Omitted category: Online and face-to-face blend.

association between synchronous tutoring—in which the interaction between the student and tutor is live or immediate—and students' math achievement. This estimated effect is substantively large and is in reference to the most expensive form (a synchronous and asynchronous combination), again suggesting no positive correlation between the hourly rates charged for different types of digital tutoring and the programs' effectiveness in increasing student achievement.[6]

These findings, combined with our analysis of student selection into different types of digital tutoring, raise potential concerns about which students have access to the relatively more effective types or forms of digital tutoring. For example, our analysis of student enrollment with digital providers showed that English language learners and students with disabilities were significantly less likely to receive OST tutoring in synchronous formats, which the value-added model estimation suggests is more effective in increasing student math achievement. In addition, students with disabilities were *more* likely to receive tutoring with a curriculum-based software program that drives student learning—which is negatively associated with student reading achievement—or via a combination of tutor-with-software driven and tutor-driven instruction that is negatively associated with math achievement. In our multisite, longitudinal study of OST tutoring, we consistently found (across sites and over time) that English language learners and students with disabilities were less likely to realize achievement gains through OST tutoring.

It is also important to reiterate, however, that given the limitations of our measures of digital tutoring characteristics and the preliminary nature of this research, we see these findings as *suggestive* of potentially troubling patterns in access to different types of digital tutoring, rather than as definitive evidence of inequitable treatment in the provision of OST tutoring. More research is needed to confirm the associations we have found among attributes of digital tutoring offerings and measures of student achievement.

TABLE 8
Value-Added With School Fixed-Effects Models of Digital Provider Effects: Instruction Driver

	Math score (standardized)		Reading score (standardized)	
Digital provider and student characteristics	Coefficient	*SE*[a]	Coefficient	*SE*
Instruction driver				
Curriculum-based software	−0.132	.084	**−0.142**	.066
Tutor-structured	−0.035	.126	−0.202	.161
Tutor with curriculum-based software	0.035	.057	−0.006	.063
Combination tutor with software-driven[b]	**−0.141**	.050	0.016	.042
Prior year standardized score	**0.334**	.067	**0.393**	.036
Attended OST tutoring last year	0.024	.033	0.034	.026
Asian	0.211	.367	0.056	.300
Hispanic	0.082	.061	0.091	.056
Other race	0.062	.190	0.098	.095
White	−0.062	.098	−0.047	.117
Free-lunch eligible	0.029	.031	**0.123**	.037
English language learner	**−0.152**	.065	−0.079	.062
Student with disability	0.018	.188	0.113	.197
Female	0.044	.038	**0.089**	.029
Percentage of days absent from regular school in prior year	**−1.669**	.760	**−2.808**	.679
Retained in grade	−0.134	.132	**−0.703**	.208
Constant	0.038	.211	**0.562**	.278

Note. Additional controls (not reported): School fixed effects and grade year. OST = out-of-school time.
[a]Robust standard errors.
[b]Omitted category: Software-driven and tutor-driven.

Furthermore, our empirical analysis of tutoring effects is limited to just one of the five sites in our larger study, and we have seen across school districts how administrative policies and practices can also influence access to quality OST tutoring and its effectiveness in increasing student achievement.

Conclusions and Implications for Policy and Future Research

Although exploratory, our study of digital OST tutoring illustrates the many dimensions along which digital tutoring may vary, including the role and location of the tutor, the type of software used, and the nature of the curriculum, as well as the extent to which these varying attributes might potentially be associated with digital providers' effectiveness in increasing student achievement. Indeed, these are not technical, peripheral variables in the instructional settings of digital tutoring, but rather, our qualitative work suggests they may matter as much as other well-established factors such as time on task, teacher qualifications, student–teacher ratio, and so on, in explaining instructional effects in traditional classrooms.

We also considered the significance of these patterns in the context of broader patterns of student characteristics and participation in OST programming overall. In our prior work, we have found that English language learners and students with disabilities are more likely to attend OST tutoring (Heinrich et al., 2014). This is good news given the intended focus of educational reform efforts on these subgroups, but it will be dampened if other research confirms our findings, suggesting that students with special needs are less likely to receive the more effective forms of digital OST tutoring.

Furthermore, our analysis suggests that digital providers are more rapidly gaining market share

TABLE 9
Value-Added With School Fixed-Effects Models of Digital Provider Effects: Curriculum Location

	Math score (standardized)		Reading score (standardized)	
Digital provider and student characteristics	Coefficient	*SE*[a]	Coefficient	*SE*
Curriculum location				
Curriculum location only digital online	−0.078	.051	−0.047	.046
Digital-online and digital-local combination[b]	**−0.159**	.053	0.001	.037
Prior year standardized score	**0.334**	.067	**0.391**	.036
Attended OST tutoring last year	0.024	.033	0.037	.026
Asian	0.213	.365	0.062	.297
Hispanic	0.080	.061	0.092	.057
Other race	0.058	.193	0.097	.093
White	−0.065	.096	−0.043	.116
Free-lunch eligible	0.028	.031	**0.122**	.037
English language learner	**−0.152**	.066	−0.078	.063
Student with disability	0.016	.188	0.111	.197
Female	0.046	.038	0.087	.029
Percentage of days absent from regular school in prior year	**−1.667**	.761	**−2.814**	.682
Retained in grade	−0.133	.131	**−0.707**	.207
Constant	0.038	.211	**0.557**	.278

Note. Additional controls (not reported): School fixed effects and grade year. OST = out-of-school time.
[a]Robust standard errors.
[b]Omitted category: Digital-online and local-nondigital combination.

than providers of face-to-face private tutoring, while they are charging higher hourly rates and delivering fewer hours of OST tutoring to students. These higher rates might be justified if students and families were getting higher quality services for their money, but our exploratory research comparing the effectiveness of digital versus nondigital providers, as well as different types of digital providers, does not find positive linkages between tutoring quality and rates charged. In addition, our longitudinal, multisite study in five large, urban districts has consistently shown a very strong association between hours of tutoring received and OST tutoring effectiveness in increasing student achievement (Heinrich et al., 2014). The significantly lower number of hours of OST tutoring received by students served by digital (vs. nondigital) tutoring providers also likely contributes to the overall negative correlation we find between digital tutoring and student mathematics and reading achievement (when compared with students served by nondigital providers).

It is also important to emphasize one more time, however, the clear need for more research to support greater understanding of the effects of particular forms of digital tutoring on student achievement and the characteristics of the instructional setting that may contribute to or hinder positive effects. In addition, further research is needed to disentangle attendance patterns and program effects by subgroups, including family socioeconomic background, with specific attention to students from low-income settings. The potential for selection bias in our quantitative analysis remains, and this type of research would also be important for improving our specification of models for estimating program effects.

Our field research also illuminates the challenges in documenting and measuring technology use and the many pathways through which it might mediate the effectiveness of educational interventions on student learning. As digital programming continues to expand, there is an urgent need for more rigorous, independent evaluations

TABLE 10

Value-Added With School Fixed-Effects Models of Digital Provider Effects: Tutor Synchronicity

Digital provider and student characteristics	Math score (standardized)		Reading score (standardized)	
	Coefficient	*SE*[a]	Coefficient	*SE*
Tutor synchronicity				
Asynchronous	−0.069	.143	−0.016	.093
Synchronous[b]	**0.104**	.037	0.032	.037
Prior year standardized score	**0.335**	.067	**0.391**	.036
Attended OST tutoring last year	0.036	.033	0.039	.026
Asian	0.198	.360	0.063	.300
Hispanic	0.092	.058	0.092	.058
Other race	0.080	.192	0.099	.094
White	−0.053	.095	−0.042	.116
Free-lunch eligible	0.027	.031	**0.124**	.037
English language learner	**−0.159**	.065	−0.076	.062
Student with disability	0.018	.187	0.110	.197
Female	0.045	.039	**0.088**	.029
Percentage of days absent from regular school in prior year	**−1.705**	.767	**−2.818**	.682
Retained in grade	−0.142	.131	**−0.710**	.207
Constant	−0.022	.204	0.553	.284

Note. Additional controls (not reported): School fixed effects and grade year. OST = out-of-school time.
[a]Robust standard errors.
[b]Omitted category: Combination of synchronous and asynchronous.

of its effectiveness to inform federal, state, and local policy decisions regarding the role and application of technology in educating underserved students. Currently, the limited, self-generated information that is disseminated by providers to parents and students does not usefully guide parent and student choices of digital providers or aid school districts in their program improvement efforts. Generating more accurate estimates of digital tutoring effects will require a more precise and comprehensive taxonomy of digital tutoring, as we have attempted to advance here.

We are currently engaging in new research that will help us to further test and refine our taxonomy of digital tutoring and supplemental instruction in day school as well as OST settings. We are also looking at different models for integrating face-to-face instruction (to varying extents) with content accessed digitally in different educational settings to better understand the role and importance of face-to-face instruction. Because of the number of dimensions on which digital education can vary in implementation, it is challenging to characterize and confirm what defines or determines effective practice. Yet this is critically important work for supporting the dissemination and scalability of effective digital educational practices. A recent review of studies focused on the potential for digital educational technology to support personalized instruction (Enyedy, 2014) found a lack of studies focused on the K–12 context, as did the Means, Toyama, Murphy, Bakia, and Jones (2010) and Means, Toyama, Murphy, and Bakia (2013) meta-analyses. Given the rapidly expanding and wide-ranging uses of digital educational technology in K–12 schools today, we need more efforts to compile the lessons learned from this type of research.

Acknowledgments

We express our most sincere thanks to the staff in the study school districts for their time and support for this project. We would also like to thank the staff at the Wisconsin Center for Education Research, University of Wisconsin-Madison, the University of Texas at Austin, and the University of Southern California for their help and support of this work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We would like to thank the funder of this research, the Institute of Education Sciences, PR/Award number: R305A090301, Education Policy, Finance and Systems Research Program, Goal 3.

Notes

1. Students eligible for out-of-school time (OST) tutoring under No Child Left Behind (NCLB) include those in public schools not making adequate yearly progress for at least 3 years who were also eligible for free or reduced-price lunch. Districts frequently also specify additional eligibility criteria, such as proficiency levels assessed via standardized tests, if the number of eligible students exceeds available resources.

2. Indicators were only included in this list if the averages come from at least 50 observation points (typically, there are two recorded per tutoring session) for digital sessions and 50 for nondigital.

3. For a fuller discussion of these findings and the research, see Good, Burch, Stewart, Acosta, and Heinrich (2014).

4. Within each dimension, we also added all combinations of classification options as classification options themselves. So, for example, we could accurately characterize a tutoring program that integrally features a combination of both Web-based and hard-copy curricula without dropping any information. We used this combinatorial option in at least two important contexts: Tutoring programs that incorporate multiple modes of service, all of which every enrolled student experiences at different times or during different sessions in the program. A program in which students independently complete curriculum-based software lessons installed on their iPods before meeting every week with in-person tutors would be an example. Tutoring programs that offer multiple modes of service, and each student chooses one of those modes at the outset of their enrollment, in effect creating multiple distinct subprograms. A program in which some students always work with their tutors in a physical classroom while other students always work with their tutors online would be an example.

5. The full set of results from these descriptive and logistic regression analyses are available upon request from the authors.

6. Across these models, approximately 16% of the variation in changes in math achievement and 35% of the variation in changes in reading achievement are explained by the models (as indicated by R^2 measures). The substantive results regarding the effects of the various provider attributes also hold when the measures of their different characteristics are combined into a single model for estimating changes in math (and reading) achievement. In addition, we estimated all of these models with student math and reading gains as the outcome (instead of controlling for prior student test scores on the right-hand side of the model) and found that the results on digital provider attributes were substantively the same.

References

Arroyo, B., Tai, W., Muldner, M., Woolf, K., & Park, B. (2013). Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational Psychology*, *105*(4), 957.

Bakia, M., Shear, L., Toyama, Y., & Lasseter, A. (2012). *Understanding the implications of online learning for educational productivity*. Washington, DC: U.S. Department of Education.

Barbour, M. K., & Reeves, T. C. (2009). The reality of virtual schools: A review of the literature. *Computers and Education*, *52*, 402–416.

Beckett, M., Borman, G., Capizzano, J., Parsley, D., Ross, S., Schirm, A., & Taylor, J. (2009). *Structuring out-of-school time to improve academic achievement: A practice guide* (NCEE #2009-012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Bingham, A. J. (in press). Drowning digitally? How disequilibrium shapes practice in a blended learning charter school. *Teachers College Record*.

Brush, T., & Hew, K. F. (2006, December). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research and Development*, pp. 224–252.

Burch, P., & Good, A. (2013). *Supplemental educational services observation instrument*. Madison: Wisconsin Center for Education Research.

Burch, P., & Good, A. (2014). *Equal scrutiny: Privatization and accountability in digital education*. Cambridge, MA: Harvard Education Press.

Cavanaugh, C., Barbour, M., & Clark, T. (2009). Research and practice in K-12 online learning: A review of open-access literature. *International Review of Research in Open and Distance Learning*, *10*, 1492–3831.

Cole, R., Kemple, J. J., & Segeritz, M. D. (2012). *Assessing the early impact of School of One: Evidence from three school-wide pilots*. New York, NY: The Research Alliance for New York City Schools.

Durlak, J., & Weissberg, R. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: CASEL.

Durlak, J., Weissberg, R., & Pachen, M. (2010). Meta-analysis of after-school programs that promote social skills in adolescents. *American Journal of Community Psychology*, *45*, 294–309.

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605–619.

Enyedy, N. (2014). *Personalized instruction: New interest, old rhetoric, limited results, and the need for a new direction for computer-mediated learning*. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/personalized-instruction

Farkas, G., & Durham, R. (2006, February). *The role of tutoring in standards-based reform*. Presented at the Standards-Based Reform Conference at University of Wisconsin-Madison.

Good, A., Burch, P., Stewart, M., Acosta, R., & Heinrich, C. (2014). Instruction matters: Lessons from a mixed-method evaluation of out-of-school time tutoring under No Child Left Behind. *Teachers College Record*, *116*(3), 1–34.

Goslee, S., & Conte, C. (1998). *Losing ground bit by bit: Low-income communities in the information age*. Washington, DC: Benton Foundation. Retrieved from www2.ctcnet.org/ctc/benton/low-income.pdf

Heinrich, C., Burch, P., Good, A., Acosta, R., Cheng, E., Dillender, M., . . . Stewart, M. (2014). Improving the implementation and effectiveness of out-of-school-time tutoring: A Longitudinal multisite, mixed-method investigation. *Journal of Policy Analysis and Management*, *33*(2), 471–494.

Heinrich, C. J., Meyer, R., & Whitten, G. (2010). Supplemental educational services under No Child Left Behind: Who signs up, and what do they gain? *Educational Evaluation and Policy Analysis*, *32*, 273–298.

Heinrich, C. J., & Nisar, H. (2013). The efficacy of private sector providers in improving public educational outcomes. *American Educational Research Journal*, *50*, 856–894.

Heistad, D. (2007). *Evaluation of supplemental education services in Minneapolis Public Schools: An application of matched sample statistical design*. Minneapolis, MN: Office of Research, Evaluation and Assessment, Minneapolis Public Schools.

Hock, M., Pulvers, K., Deshler, D., & Schumaker, J. (2001). The Effects of an after-school tutoring program on the academic performance of at-risk students and students with LD. *Emotional and Special Education*, *22*, 177–186.

iNACOL. (2011). *The online learning definitions project*. Vienna, VA: International Association for K-12 Online Learning. Available from www.incaol.org

Lauer, P., Akiba, M., Wilkerson, S., Apthorp, H., Snow, D., & Martin-Glenn, M. (2006). Out-of-School-Time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, *76*, 275–313.

Little, P., Wimer, C., & Weiss, H. (2008). *Programs in the 21st century: Their potential and what it takes to achieve it*. Cambridge, MA: Harvard Family Research Project.

London, R., Pastor, M., & Rosner, R. (2008, Spring). *When the divide isn't just digital: How technology-enriched afterschool programs help immigrant youth find a voice, a place and a future*. Chicago, IL: Afterschool Matters.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*, 423–458.

Margolin, J., Kieldon, B., Williams, R., & Schmidt, M. C. (2011). *Vermont's Title II-D enhancing education through technology program: 2010-2011 final report*. Los Angeles, CA: American Institute for Research.

Means, B., Bakia, M., & Murphy, R. (2014). *Learning online: What research tells us about whether, when and how*. New York, NY: Routledge.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service.

Means, B., Toyama, Y., Murphy, R. F., & Bakia, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*, 1–47.

Miron, G., & Urschel, J. L. (2012). *Understanding and improving full-time virtual schools*. Boulder, CO: National Education Policy Center.

Molnar, M., & Cavanagh, C. (2014, January 22). K-12 Publishing, Ed-Tech Markets Seeing Sales Increases. *Education Week*, *33*(18), 9.

Price, L., Richardson, J., & Jelfs, A. (2007). Face-to-face versus online tutoring support in distance education. *Studies in Higher Education*, *32*(1), 1–20.

Queen, B., Lewis, L., & Coopersmith, J. (2011). *Distance education courses for public education elementary and secondary students 2009-2011*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2012/2012008.pdf

Richards, J., & Struminger, R. (2013). *2013 U.S. education technology industry market: PreK-12*. Washington, DC: Software & Information Industry Association.

Rickles, J. H., & Barnhart, M. K. (2007). *The impact of supplemental educational services participation on student achievement: 2005-06* (Report, Pub. No. 352). Los Angeles, CA: Los Angeles Unified School District Program Evaluation and Research Branch, Planning, Assessment and Research Division.

Rose, R. M., & Blomeyer, R. L. (2007). *Access and equity in online classes and virtual schools*. North American Council for Online Learning. Available from www.inacol.org

Slattery, C. (2003). *The Impact of a computer-based training system on strengthening phonemic awareness and increasing reading ability level* (Doctoral dissertation). Widener University, Chester, PA.

Snyder, T. D., & Dillow, S. A. (2013). *Digest of Education Statistics 2012* (NCES 2014-015). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Springer, M. G., Pepper, M. J., & Ghosh-Dastidar, B. (2009). *Supplemental educational services and student test score gains: Evidence from a large, urban school district* (Working paper). Nashville, TN: Vanderbilt University.

Stevens, P. (2012). *The William T. Grant Foundation: Contributions to the after-school field*. New York, NY: W.T. Grant Foundation. Available from http://www.wtgrantfoundation.org

Vandell, D., Reisner, E., & Pierce, K. (2007). *Outcomes linked to high-quality afterschool programs: Longitudinal findings from the study of promising practices*. Irvine: University of California.

Zhao, Y., Lei, J., Yan, B., Tan, H. S., & Lai, C. (2005). What makes the difference? A practical analysis of effectiveness of distance education. *Teachers College Record*, *107*, 1836–1884.

Zickuhr, K., & Smith, A. (2012). *Digital differences*. Pew Research Center. Available from www.pewinternet.org

Zimmer, R., Gill, B., Razquin, P., Booker, K., & Lockwood, J. R. (2007). *State and local implementation of the No Child Left Behind Act: Volume I-Title I school choice, supplemental educational services, and student achievement*. Washington, DC: RAND.

Zimmer, R., Hamilton, L., & Christina, R. (2010). After-school tutoring in the context of No Child Left Behind: Effectiveness of two programs in the Pittsburgh Public Schools (PPS). *Economics of Education Review*, *29*(1), 18–28.

Authors

PATRICIA BURCH is an associate professor at the Rossier School of Education. Her work focuses on issues of equity and access in privatization trends in K–12 education. She is the author of *Hidden Markets: The New Education Privatization*.

ANNALEE GOOD is an associate researcher at the Wisconsin Center for Education Research at the University of Wisconsin-Madison. Her research examines dynamics within the K–12 education context, including the use of digital tools, private contracting, and academic tutoring.

CAROLYN HEINRICH is the Sid Richardson Professor of Public Affairs, affiliated professor of economics, and the director of the Center for Health and Social Policy at the University of Texas at Austin. Her research focuses on education, social welfare policy, workforce development, program evaluation, and public management and performance management.

Manuscript received July 30, 2014
First revision received December 13, 2014
Second revision received March 24, 2015
Third revision received April 9, 2015
Accepted April 10, 2015