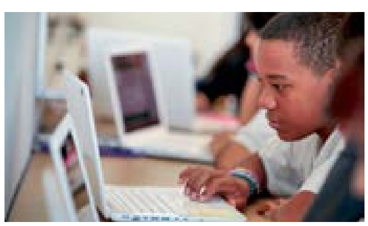
Michael & Susan Dell Foundation

IN PARTNERSHIP WITH SRI International



Robert Murphy · Eric Snow · Jessica Mislevy · Larry Gallagher · Andrew Krumm · Xin Wei

Blended Learning Report

May 2014

Michael & Susan Dell
FOUNDATION

TABLE OF CONTENTS

Michael & Susan Dell
FOUNDATION

Executive Summary

BACKGROUND

With the prevalence of digital instructional resources and access to high-speed bandwidth and devices, many districts, charter school networks, and schools and teachers are experimenting with ways to combine online instruction with regular classroom instruction to support teaching and improve students' learning experiences. This blending of online and face-to-face instruction is expected to be standard practice in many classrooms in the future.

Since blended learning is an emerging field there are currently many different conceptualizations of what it means to "do" blended learning. The definition adopted for this research report, following Staker and Horn (2012), has the following components:

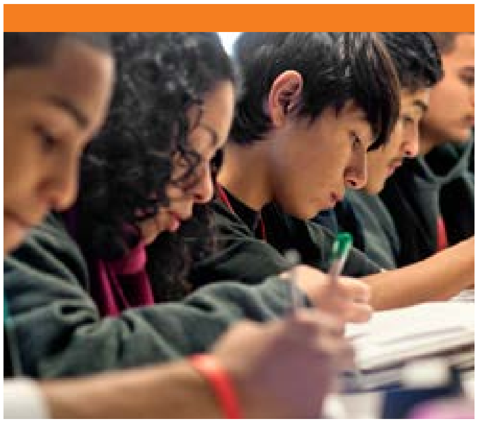
- It involves teaching and learning within a formal education program
- Students learn at least in part through online delivery of content and instruction
- Students have some level of control over time, place, path, and/or pace of instruction
- Part or all of instruction is delivered away from home in a supervised, brick-and-mortar location

With funding from the Michael & Susan Dell Foundation, the SRI International Center for Technology in Learning studied the adoption of blended learning models in selected schools in California and Louisiana. This research report presents the findings of this formative and summative research effort. During the 2011–12 school year, five charter management organizations (CMOs) received funding from the foundation to help support the development and expansion of blended learning models in a total of 13 schools serving low-income communities and families. The CMOs were Rocketship Education, Summit Public Schools, Alliance-College Ready High Schools, KIPP LA, and FirstLine Schools.

The study was motivated by the following research questions:

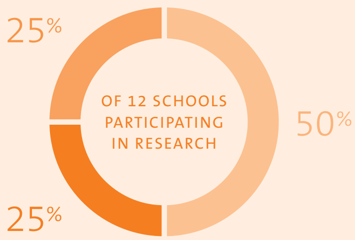
- Do students in blended learning models show changes in academic achievement that differ significantly from their peers'?
- Do students in blended learning models show a propensity to close the achievement gap?
- Are differences in the way that blended learning models are implemented associated with differences in learning outcomes?
- Are blended learning models more effective for some types of students or subject areas than for others?
- To what extent is each of the blended learning models being implemented as intended?

For many of the participating sites, the findings reported here are based on the experiences of schools in their first or second year of implementing blended learning, and five schools were in their first



WITH SUPPORT FROM:

2011–12 BLENDED LEARNING ADOPTION



FIRST YEAR OF ADOPTION
SECOND YEAR OF ADOPTION
THIRD YEAR OR MORE OF ADOPTION

TABLE ES-1. RESEARCH SAMPLE SUMMARY

CHARTER MANAGEMENT ORGANIZATION	SCHOOLS PARTICIPATING	GRADE LEVELS IN STUDY	TEACHERS PARTICIPATING	STUDENTS PARTICIPATING
Alliance	3 high schools	9, 10	26	900
FirstLine	1 elementary school	K–8	20	460
Rocketship	5 elementary schools	K–5	70	2,381
KIPP LA	1 elementary school	K–1	8	230
Summit	2 high schools	9	3	230

year of operation. Of the 12 schools participating in the research, 6 were in their first year of adoption during the 2011–12 school year and three were in their second year. At all the sites—including Rocketship, which opened its first school in 2007—the organizations and schools continue to experiment and refine their blended learning models to improve the experience and results for teachers and students. In several cases, the models studied during the 2011–12 school year changed significantly from the 2012–13 school year as the organizations and schools reflected on their results, upgraded technology, identified more promising instructional software and practices, and reallocated resources to better serve their students. Even though the models studied and the technology that supports them are evolving, our research findings are relevant to other organizations and schools that are preparing for or piloting their own blended learning models.

The sample participating in the research within each site included both elementary and high schools spanning grades K–10, all of which served students from low-income communities (Table ES-1).

BLENDED LEARNING MODELS

The models that were the subject of SRI's research are sometimes referred to as "rotation" blended learning models (Staker and Horn, 2012). During the school day, students move between regular classroom instruction and online instruction based on a schedule or at the teacher's discretion. The online instruction may be organized as one of several stations in a classroom that students rotate among during a class period. In another version, online instruction occurs separately from the core teacher-led instruction, often in a computer or learning lab. Both of these rotation blended learning models—classroom and lab based—were adopted to some degree by the schools in the research.

EVIDENCE OF THE EFFECTIVENESS OF BLENDED LEARNING

Up to this point, limited rigorous research has been conducted on the effectiveness of blended learning models, particularly in K–12 school settings (Means et al., 2009). Claims are made about the relative effectiveness of various blended learning models relative to more traditional forms of instruction, but thus far little evidence has been collected to back these claims. Clearly, more rigorous research is needed to understand the utility of blended learning to support teaching and learning for all students within K-12 learning environments.

SUMMARY OF FINDINGS FROM THE 2011–12 SCHOOL YEAR RESEARCH

This cross-site summary of select implementation findings is based on interviews, observations, and teacher surveys within each of the sites during the 2011–12 school year. The findings reported are related to issues with infrastructure and technology, elements of the software design, perceived benefits to teaching and learning, and student productivity.[1]

Infrastructure and Technology

Unreliable Internet connectivity, inadequate bandwidth, and problems with software programs hindered many schools' ability to implement their models. With a few exceptions, a majority of

[1] *Due to the variation in the blended models adopted by each research site, findings related to the impact of the various blended learning models on student learning and the schools' ability to close the high-income/low-income achievement gap are reported in each site's research profile.*

Michael & Susan Dell
FOUNDATION

the teachers at each site reported that either they or their students experienced technical problems during the year. For teachers at three of the five research sites, these problems were reported to be either moderate or significant barriers to effective use of their blended learning models. The most commonly cited technical issues were the reliability of Internet connectivity and online programs and insufficient bandwidth to run some of the programs.

> *"In general, sites were interested in identifying programs that (1) were comprehensive and aligned with the Common Core standards, (2) were adaptive, (3) were interactive and engaging, (4) were assignable, (5) provided valid reports of student learning, (6) allowed for integration into the school learning management systems and had a single sign-on portal when more than one program was used, (7) were cloud based, and (8) were affordable."*

On-site IT support and contingent instructional activities provided critical support in the wake of problems with technology. Technology-related problems need to be anticipated and planned for. Three of the five research sites hired staff to provide on-site IT support, which freed teachers from this burden. Yet even with on-site IT staff available, administrators, teachers, and lab monitors in several sites expressed the need for backup activities and software programs for the times technology issues arose. Examples of such activities in lab rotation models were the use of worksheets or off-line projects or the use of software programs that resided on the students' computers or schools' servers.

Site-based blended learning coordinators and teams provided critical technical support for administrators and teachers during the transition to blended learning. These positions were essential to reduce some of the burdens on administrators and teachers so they could stay focused on teaching and learning. People in these coordination roles were often responsible for identifying and piloting online programs, negotiating with vendors, providing reports on student performance in the online programs, coordinating professional development, and overseeing the purchase of technology.

Software Design Elements

During the school year, sites continued to experiment with different online instructional software in ELA and math. They continually refined their models to best serve teachers and students, which included searching for and piloting different online programs. In general, sites were interested in identifying programs that (1) were comprehensive and aligned with the Common Core standards, (2) were adaptive, (3) were interactive and engaging, (4) were assignable, (5) provided valid reports of student learning, (6) allowed for integration into the school learning management systems and had a single sign-on portal when more than one program was used, (7) were cloud based, and (8) were affordable.

How assignable online instruction programs are may limit their integration with the classroom curriculum. Many sites used online instructional programs such as DreamBox and Istation that are adaptive and enable students to work at their own pace on content that is appropriately challenging. Given this self-pacing, students may not be working on the same topic at the same time or at the same time the teacher is presenting the material in the classroom. Many of the programs do not allow teachers or administrators to assign the online content for students. The importance of program assignability to the implementation of a blended learning model depends on the role of online instruction in the model. That is, assignability is less critical if the online programs are being used for remediation by allowing students to proceed through the content at their own pace and level as they fill in gaps in their knowledge as needed. Assignability is more critical when teachers are using the online instruction as an integrated supplemental activity to directly support the lessons they are presenting in the classroom.

Michael & Susan Dell
FOUNDATION

Teachers in the study reported that the quality of the online software programs varied by subject. Across sites, teachers and administrators tended to be more satisfied with the content and quality of software programs developed for math skills than for English language arts (ELA). For ELA, the greatest quality concerns were about supporting writing instruction.

Several factors limited teachers' use of data from online systems to inform instruction. All of the sites participating in the research had strong cultures of using student performance data to make informed instructional decisions. In general, however, schools and teachers underutilized data from the online instructional programs. Teachers and administrators noted several challenges to leveraging the student data. One challenge related to the quality of dashboards used to report student progress in some of the programs. Teachers reported that either the dashboards were difficult to access or the information on student progress was difficult to interpret in terms of the schools' own content standards. In turn, that hindered teachers and lab monitors in efficiently identifying students who were struggling or falling behind progress expectations. Teachers also reported that progress and performance measures were often not aligned with state content standards or the schools' curriculum, leaving them less able to anticipate how students might perform on state-mandated assessments or internal benchmark exams.

Teachers' and administrators' trust in the assessments embedded in the online programs was limited; several reported they trusted their own formative assessments more than a program's measure of content mastery because of uncertainties in the validity of the program's measures.

Several sites that used more than one online program in their blended learning model faced the challenge of integrating multiple streams of student performance data from the programs into their own student information systems. Because the data from the online programs were not well integrated into a single dashboard, many teachers found that reviewing the data was too time consuming to do on a regular basis. Third-party providers such as Education Elements have entered the market to help schools package data for use by educators. (Education Elements worked with KIPP LA and Alliance during the 2011–12 school year.) However, during this study, the research sites were left to make sense of the multistream data on their own.

Benefits to Teaching and Learning

The sites use blended learning models to personalize learning through self-paced programs, adaptive online instructional content, and facilitation of small-group instruction for students with the greatest academic needs. Administrators and teachers described several uses and benefits of blended learning to personalize students' work. Self-paced and adaptive math software programs like DreamBox Learning, ST Math, and Istation support personalized learning by individualizing instruction, challenging students with the appropriate level of content, and directing their path through the program based on prior successes. Students progress through the material at different speeds based on their learning needs. They may take more time to complete a given topic, skip topics they already know, repeat topics they need help on, and explore content that is beyond the curriculum. A majority of the teachers surveyed reported that technology and computer-mediated instruction played a major role in providing enrichment for advanced students and remediation for struggling students.

Learning was further personalized within all the sites by teachers using blended learning to facilitate small group instruction. For example, students with the greatest academic needs at FirstLine's Arthur Ashe school spent their lab time in a "learning support" classroom, rotating in small groups between teacher-led activities, online workstations, and independent practice. In KIPP Empower Academy's in-class rotation model, students during English language instruction were grouped by ability and teachers differentiated their instruction based on the level of each group. As one group of students worked on adaptive online software programs, other students received instruction from teachers in the classroom.

Teachers report that blended learning benefits students' procedural skills development more than higher order thinking. Perhaps because of the online instructional programs' emphasis on providing students with opportunities to practice basic skills and procedures, a majority of teachers surveyed across sites reported that their blended models were more effective at helping students recall basic facts than at helping them develop higher order thinking skills. This finding held true across subject areas. Teachers' perceptions of student skill development reflected their schools' blended learning model as a whole— the combination of online and offline instruction—and not just the online component.

"According to a majority of the administrators, teachers, and lab monitors interviewed, weekly goal-setting helped students to become more invested in their learning and to see both the rewards of meeting goals and consequences of failing to meet them."

Goal-setting promotes self-directed learning in the blended learning environment. The goal-setting practices sites used to direct student progress in the blended instructional environment tended to promote within students a sense of accountability and ownership over the learning process. Most often, goal setting was associated with students' weekly goals for progress on the online programs. According to a majority of the administrators, teachers, and lab monitors interviewed, weekly goal-setting helped students to become more invested in their learning and to see both the rewards of meeting goals and consequences of failing to meet them.

Teachers reported that students' readiness for self-directed learning may vary by their academic preparation. While teachers cited advantages with their school's blended learning model, they generally indicated that students did not reap these benefits equally. Across sites, a large majority of teachers reported that their models were at least somewhat effective for all students but those who benefited most were students whose academic work was either ahead of, or at the same level as, most students their age. The models were possibly less effective for students whose academic work was below their expected grade level. In addition, many teachers interviewed felt a student's ability to self-manage and self-direct their learning determined which children would most likely thrive in a blended learning model.

Teachers' satisfaction with their school's blended learning models varied. Most teachers at FirstLine, KIPP Empower Academy, and Summit reported that they were satisfied with blended learning in their instruction and would recommend it to other teachers. However, approximately half the teachers at Alliance and Rocketship reported they were less satisfied. Data collected during site visits and responses to the teacher survey suggest some explanations for these differences. Within Rocketship schools, teachers by design were largely disconnected from what took place within the learning lab, and this may have influenced their feelings about the online instructional component of the model. Many Alliance teachers spoke of the classroom management challenges in implementing an in-class rotation model with large class sizes and high school students. Other teachers at Alliance, particularly those in subjects other than math, also reported that the three-station model (direct instruction, independent, and collaboration) was not ideal for the content they were teaching.

Student Productivity

Strong classroom/learning lab management practices are critical to ensure student productivity in online environments. Teachers and lab monitors alike stressed the importance of establishing the proper academic culture, norms, and behavior management practices for a blended learning model to be successful. This is especially important in managing student transitions between online and teacher-led instructional blocks, particularly with younger children.

Implications and Future Research

Charter management organizations and schools in the study during the 2011–12 school year experimented with ways of organizing their instruction to best meet the needs of their students. Throughout the study, many of the sites continued to pilot different online offerings and ways of combining and integrating online and face-to-face instruction. All sites used the online instructional component of their model to supplement their core instruction, providing students additional opportunities to practice skills they had just learned for remediation and enrichment. None of the schools, with the exception of one Alliance high school, used online learning programs as the primary source of instruction.

Given the schools' early-stage adoption of blended learning and the continuing evolution of the models at each site, it is much too early to judge the effectiveness of these models relative to other ways to organize instruction. As new and refined blended learning models emerge and show promise, more rigorous evaluations are needed to study their benefits at scale and identify the design and features needed to help all teachers and students succeed in a blended learning environment.

During this research, several important facilitating factors emerged that will most likely deserve consideration by others looking to adopt blended learning models like those implemented by schools in the study:

Blended learning coordinators played an important role in supporting schools' adoption of blended learning. Each of the research sites had someone or a team of people who supported the schools in their adoption of blended learning, freeing administrators and teachers of the logistical burdens associated with the adoption process including negotiating with software and hardware vendors, consultants, and Internet providers. Without such support, these schools probably would not have had the success in implementing their models that they did this early in their adoption cycle.

Establishing productive, self-directed learning cultures is important for students to fully benefit from online learning. Students as young as 5 and 6 years old are spending up to a few hours each day receiving online instruction in a number of schools in the study, including those associated with Rocketship, FirstLine's Arthur Ashe Academy, and KIPP Empower Academy. Each of these sites spent considerable time over the prior summer and throughout the school year helping students make efficient transitions between regular classroom instruction and their online instruction time. Setting weekly goals for student progress as a mechanism to focus students and increase productivity was a practice that emerged across the school year in most of the schools in the study.

Single sign-on portals allow even very young children to quickly access online programs. Reliable, single sign-on portals were an important feature of efficient transitions between teacher-led and online instructional activities.

A reliable Internet connection and sufficient bandwidth are critical. A majority of teachers and staff managing learning labs reported that insufficient bandwidth and unreliable Internet connections were significant barriers to their effective adoption of blended learning.

On-site IT support and backup plans are critical to buffer schools from the inevitable technology issues. Schools implementing blended learning must anticipate and plan for technology-related problems that interfere with the delivery of instruction. Across all schools, teachers reported that technology issues had greatly reduced the schools' ability to effectively implement their models.

Further Research and Development

As more schools plan to adopt blended learning models, research and development efforts should consider the support teachers and students need to more effectively benefit from blended learning environments.

Online programs must be designed to support integration of content into the curriculum. For schools and teachers to better utilize online content in their instruction, it needs to be designed and organized so it can be flexibly integrated with regular teacher-led instruction. At a minimum, the content needs to be aligned with the school's curriculum standards or allow teachers or curriculum coordinators to easily map the content to their local standards. The online content must be searchable and assignable and clearly mapped to grade-level state content standards.

Reports of student progress within the online programs must be easy for administrators and teachers to access and interpret. Developers and system integrators need to address several challenges to realize the promise of online learning programs as a tool to collect and report real-time learning data to students and teachers. First, administrators and teachers must be able to trust that online programs' measures of content mastery (when available) are valid and can be used to inform instruction and identify students who are struggling and need more support. For a majority of the online programs used by schools in this study, this was not the case. To monitor student learning, most administrators and teachers continued to rely on their own formative assessments. Underutilization of the online data is compounded when multiple online programs are being used, each with its own way of measuring and communicating student performance and progress. The burden for administrators and teachers of interpreting and making sense of these multiple data streams for instructional purposes is too great, so potentially valuable real-time learning data are not being used. Until there is standardization of what constitutes mastery across online programs providing instruction in the same subject area, the potential benefits of having ready access to real-time learning data will most likely go untapped.

Finally, we need a better understanding of how to help all students succeed in self-directed online learning environments. Some students will likely be better prepared to succeed in online learning environments than others. Research is needed to understand the role of noncognitive student characteristics—e.g., motivation, persistence, resourcefulness—in students' success in self-directed online learning environments and the features of online learning programs that better support learners with different levels of preparation. More emphasis is needed to understand who is flourishing and failing in these environments and why. Programs and blended learning models can then be developed with the supports that help all students succeed.

Michael & Susan Dell FOUNDATION

Introduction

With the prevalence of digital instructional resources and access to high-speed bandwidth and devices, many districts, charter school networks, and schools are experimenting with ways of combining online and regular classroom instruction to support teaching and improve students' learning experiences. This blending of online and face-to-face instruction is expected to be standard practice in many classrooms in the future. Some districts and charter management organizations (CMOs) are already creating new administrative positions and teams combining expertise in instruction, technology, and subject area content to help their schools investigate appropriate digital resources, negotiate with vendors, and plan for technology upgrades and professional development. Also, web-based services and knowledge-sharing networks are beginning to emerge, like Gooru and Learnist that help teachers efficiently identify and assemble curated online resources to support their classroom instruction.

Blended learning models are being adopted for a variety of purposes. Some are extending instructional time at lower costs by using online programs to deliver instruction during an extra block of class time while paraprofessionals monitor the classroom or lab. Teachers are using blended models to add variety to their instruction and to free them to spend more time working with individual students. Schools are also using the adaptive and self-paced nature of some online systems to reengage students, particularly those with the greatest academic needs, by enabling them to learn at their own pace and appropriate level of challenge.

During the 2011–12 school year, five charter management organizations received funding from the Michael & Susan Dell Foundation to help support the development and expansion of blended learning models in a total of 12 schools serving low-income communities and families in California and Louisiana. The CMOs were Rocketship Education, Summit Public Schools, Alliance-College Ready High Schools, KIPP LA, and FirstLine Schools. In addition to supporting these schools, the Dell family foundation contracted with FSG to develop case studies on the CMOs and their models and with SRI International to conduct formative and summative research and investigate the potential for these models to improve student learning and close the achievement gap between schools serving low-income and high-income communities. This research report presents the findings of SRI's research. FSG's case studies can be found on the foundation's website (www.msdf.org/blendedlearning).

BLENDED LEARNING MODELS

Recent reports by the Clayton Christensen Institute for Disruptive Innovation highlight these trends toward blended learning and provide a taxonomy to describe the types of models that are emerging (Staker and Horn, 2012). Models range from brick-and-mortar schools where instruction in core subject areas is completely online to models where online instruction is used in English language arts (ELA) and math classrooms to provide additional opportunities to practice and develop skills while the teacher provides primary instruction.

> *"This blending of online and face-to-face instruction is expected to be standard practice in many classrooms in the future."*

Since blended learning is an emerging field there are currently many different conceptualizations of what it means to "do" blended learning. The definition adopted for this research report, following Staker and Horn (2012), has the following components:

- It involves teaching and learning within a formal education program
- Students learn at least in part through online delivery of content and instruction
- Students have some level of control over time, place, path, and/or pace of instruction
- Part or all of instruction is delivered away from home in a supervised, brick-and-mortar location

The models that were the subject of SRI's research are sometimes referred to as "rotation" blended learning models (Staker and Horn, 2012). During the school day, students move between regular classroom instruction and online instruction based on a schedule or at the teacher's discretion. The online instruction may be organized

TABLE 1. BLENDED LEARNING ROTATION MODEL TYPE, BY RESEARCH SITE

CLASSROOM BASED	LAB BASED
Alliance	FirstLine
KIPP LA	Rocketship
Summit	

as one of several stations in a classroom that students rotate among during a class period. In another version, online instruction occurs separately from the core teacher-led instruction, often in a computer or learning lab. Both of these rotation blended learning models—classroom and lab based—were adopted to some degree by the schools in the research (Table 1). Descriptions of the models adopted in school year 2011–12 are in the individual site profiles in this report and in the case studies on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).

> *"Clearly, more rigorous research, like the current study sponsored by the Michael & Susan Dell Foundation, is needed to understand the utility of blended learning to support teaching and learning for K–12 students."*

EVIDENCE OF THE EFFECTIVENESS OF BLENDED LEARNING

Claims are made about the relative effectiveness of various blended learning models relative to more traditional forms of instruction, but thus far little rigorous research evidence has been collected to back these claims, particularly in K–12 school settings. A synthesis of the research on online instruction effectiveness that SRI conducted with funding from the U.S. Department of Education (Means et al., 2009) is often cited by others to justify investments in blended learning in K–12 classrooms. Yet the authors found only 46 studies between 1996 and 2008 that had designs that could provide reliable evidence on the effectiveness of online instruction (20 of these studies involved blended learning models). Of those studies, only 5 studies were in K–12 settings, all of which used a blended learning model to teach a range of subjects including Spanish, Algebra I, writing, science, and social science.

In the 20 blended learning studies, the majority in higher education and medical training, the general findings were promising. Students who were exposed to blended learning did significantly better on assessments of learning than students receiving face-to-face instruction only and students receiving online instruction only. The SRI authors were careful to note, however, that the blended learning study conditions often included additional learning time and instructional elements students in control conditions did not receive. This suggests that the positive effects associated with blended learning cannot be attributed to the addition of the online component alone or to the blending of online and teacher-led instruction. In addition, the SRI authors cautioned against generalizing these findings to K–12 learning environments because so few studies in the synthesis involved K–12 students. Clearly, more rigorous research, like the current study sponsored by the Michael & Susan Dell Foundation, is needed to understand the utility of blended learning to support teaching and learning for K–12 students.

Michael & Susan Dell FOUNDATION

This Research

For many of the sites participating in this research, the findings reported here are based on the experiences of schools in their first or second year of implementing blended learning, and five schools were in their first year of operation. Of the 12 schools participating in the research, 6 were in their first year of adoption during the 2011–12 school year and 3 were in their second year. In all the sites—including Rocketship, which opened its first school in 2007— the organizations and schools continue to experiment and refine their blended learning models to improve the experience and results for teachers and students. In several cases, the models studied during the 2011–12 school year changed significantly for the 2012–13 school year as the schools reflected on their results, upgraded technology, identified more promising instructional software and practices, and reallocated resources to better serve their students. Even though the models studied and the technology that supports them are evolving, our research findings are relevant to other organizations and schools that are preparing for or piloting their own blended learning models.

The primary research questions were the following:

- Do students in blended learning models show changes in academic achievement that differ significantly from their peers'?
- Do students in blended learning models show a propensity to close the achievement gap?
- Are differences in the way that blended learning models are implemented associated with differences in learning outcomes?
- Are blended learning models more effective for some types of students or subject areas than for others?
- To what extent is each of the blended learning models being implemented as intended?

2011–12 BLENDED LEARNING ADOPTION

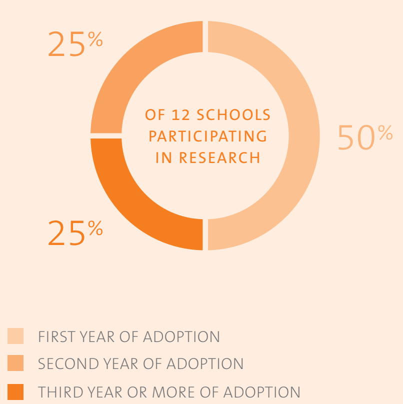
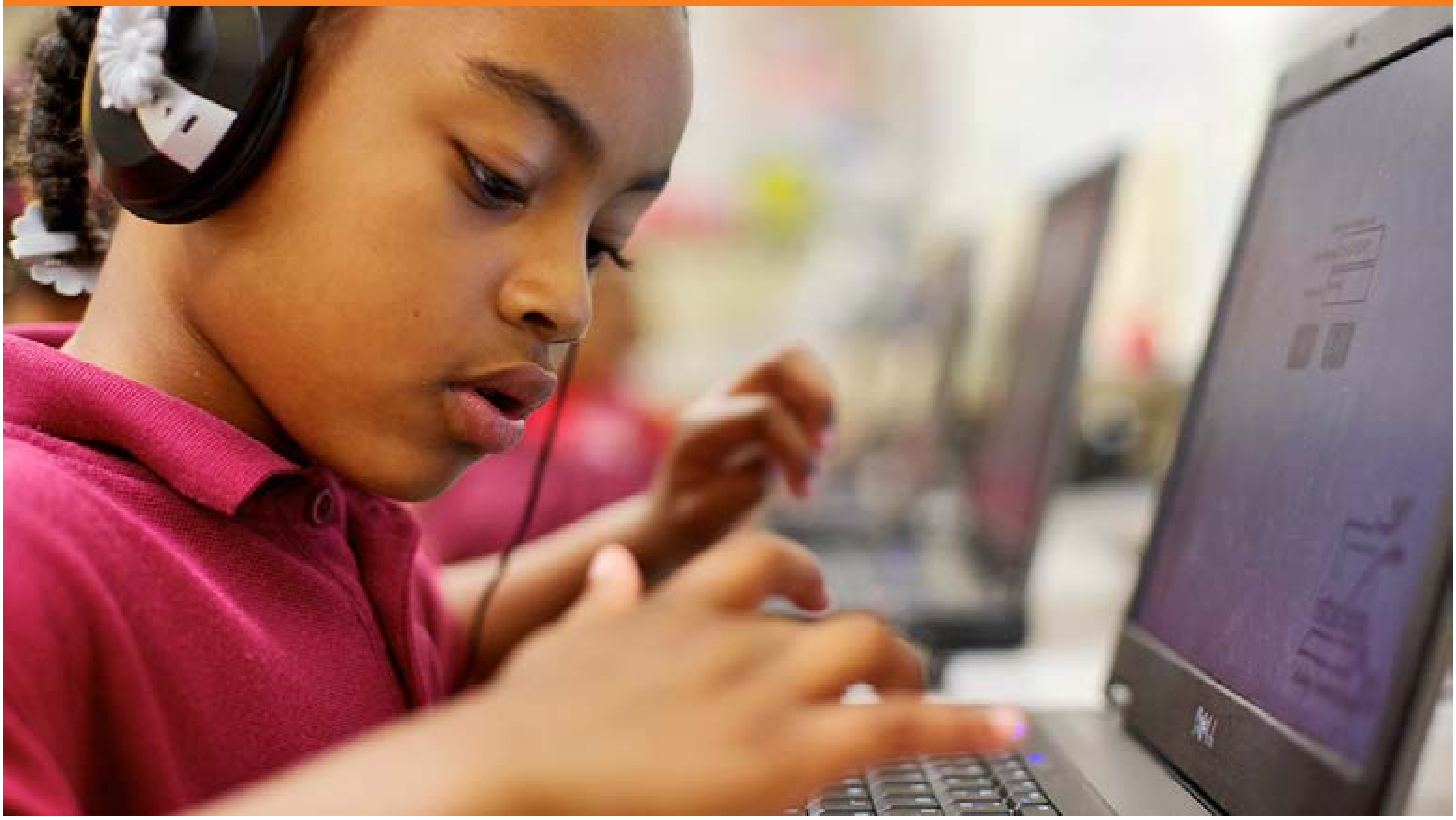


TABLE 2. RESEARCH SAMPLE SUMMARY

RESEARCH SITE	PARTICIPATING SCHOOLS	GRADE LEVELS IN STUDY	TEACHERS PARTICIPATING	STUDENTS PARTICIPATING
Alliance	3 high schools	9, 10	26	900
FirstLine	1 elementary school	K–8	30	460
Rocketship	5 elementary schools	K–5	70	2,381
KIPP LA	1 elementary school	K–1	8	230
Summit	2 high schools	9	3	220



Sample

The sites participating in the research included both elementary and high schools spanning grades K–10 (Table 2), all of which served students from low-income communities. A majority of the schools—but not all—were considered moderate to high performing based on their states' accountability rankings (see individual site profiles for details on state rankings).

Data Collection

SITE VISITS

Members of the research team visited each site in October 2011 and again in spring 2012. The visits were typically conducted over 1 or 2 days. Interviews with site and school leadership, participating teachers, and—in three sites (Alliance, Summit, and FirstLine)—students. Observations were also recorded in the respective classrooms and computer labs.

The research team drafted site visit protocols for leadership, administrator, teacher, and student interviews and classroom observations. Interview protocols covered the following topics as applicable:

- leadership's vision for blended learning, blended model description and adoption status
- changes in student and teacher roles as a result of blended learning
- review and use of system data
- perceptions of software quality supports and infrastructure required for implementation
- benefits and challenges of blended learning
- monitoring and evaluating the model's success
- reflections and lessons learned
- plans for future use and model refinement

The low-inference observation protocols were used to collect information on the following topics:

- organization of the classroom or lab
- instructional activity flow and time allocation
- student and teacher use of blended resources
- incorporation of online or computer-mediated activities
- alignment and integration of blended resources with the curriculum
- interactions among teachers and students
- student engagement while using the online programs
- access to technology

Further details about the site visits are included in the individual site profiles.

TEACHER SURVEY

In spring 2012, SRI conducted surveys of teachers implementing blended learning at participating sites and, when available, comparison teachers at other network schools not implementing blended learning (Alliance and Summit). The surveys covered the following topics, as applicable:

- organization of classroom
- access to technology
- time allocated to various instructional activities
- role of technology in supporting instruction
- review and use of system data
- satisfaction with blended learning and its effectiveness
- training participation and satisfaction
- potential technology and nontechnology barriers to effective use
- respondent background information

In addition, for Rocketship and FirstLine, we surveyed the nonteaching staff in each school responsible for managing the learning labs. Survey response rates for each of the sites are reported in the individual profiles.

SYSTEM LOG DATA

SRI worked with each research site and—in some cases the online providers—to access the automated teacher and student use data archived by the primary online instructional programs. The type and quality of the data varied greatly by online program and some of the data were not adequate for analysis purposes. With the data collected, SRI described use of the primary online programs in each site and how use varied both within and across schools when multiple schools within a site were participating. Data associated with secondary uses of technology, such as the students' use of Internet-based search tools for research and production tools like Microsoft Word and Microsoft Excel, were not captured and analyzed, nor was use of apps on tablet devices like the iPad. Data from these applications were not readily available for research purposes. SRI also conducted exploratory analyses of the relationships between students' prior achievement and intensity of use indicators. For Summit and the KIPP Empower Academy, SRI also examined the relationship between frequency of use and current student achievement. Details of the site-specific system log data collected and analyzed are in the individual site profiles.

SECONDARY DATA COLLECTION

SRI negotiated with each of the research sites for access to student demographic and achievement data. Whenever possible, available current and prior achievement data was collected—going back to the 2008–09 school year, including state test scores, other standardized test scores, and internal benchmark assessments if administered.

OUTCOME MEASURES

For the analyses of impacts on student learning, SRI included several different achievement measures in the subject areas of math and ELA. The learning outcome measures included state assessments, subject-specific end-of-course exams, high school exit exams, internal CMO benchmark assessments, and third-party assessments (e.g., TerraNova, Northwest Evaluation Association's Measurement of Academic Progress). Whenever possible, multiple measures were analyzed to determine whether the impact evidence was consistent across the different assessments. Students' scores from the prior year were also used as a measure of prior achievement and were included as a variable in the analysis models to adjust for any differences in prior achievement between students in the schools using blended learning models and the students in the comparison group. For a particular outcome measure to be included in the impact analyses, 75% or more of the eligible students in a school and grade level of interest had to have "complete" data—scores on both the outcome and prior achievement measure.

ANALYTICAL MODELS FOR ESTIMATING IMPACTS

Quasi-experiments were used to collect evidence of the potential effects of each of the sites' blended learning models on student learning outcomes. The specific design applied in each of the sites was determined by a number of factors including the adoption plans for the sites (e.g., whether they planned to roll out the model in all schools in the network during school year 2011–12) and the quality and availability of student-level data.

Key to the designs used was the identification of a valid counterfactual or comparison group that represented what would have happened to the students in schools adopting a blended learning model if they had not been in those schools. Although random assignment to experimental condition is the ideal way to construct the counterfactual and is considered the gold standard of causal research designs, this design was not feasible in any of our research sites in this report. As a result, two different quasi-experiments were conducted. Outcome measures for students in the blended learning schools during the 2011–12 school year were compared with achievement outcomes of one or both of the following comparison groups:

1. Students from the 2011–12 cohort in other schools in the CMO's network not implementing blended learning
2. A virtual comparison group created by the Northwest Evaluation Association (NWEA) using scores from the Measures of Academic Progress (MAP).

The two designs are described below. Additional details as well as the analytical models used are in the individual site profiles and the Technical Appendix.[2]

COMPARISON WITH OTHER SCHOOLS IN THE CHARTER MANAGEMENT ORGANIZATION NOT IMPLEMENTING BLENDED LEARNING

Summit, FirstLine, and Alliance all were piloting blended learning models during the 2011–12 school year in a portion of their schools, so other schools in the network were available for a comparison condition.[3] Scores for spring learning outcome assessments for students in the blended learning and non-blended learning schools were compared after controlling for prior achievement. Before conducting the analyses, SRI tested for any differences in prior achievement scores between students in the blended learning schools and those in the comparison schools.

Although the comparison schools may be similar in many ways to the schools implementing blended learning, important differences may still exist between the two group of schools (e.g., differences in curriculum, academic culture, school leadership, results, and teaching staff) that may explain differences between the schools on spring test scores or outcomes that are unrelated to the implementation of blended learning. Thus, this design cannot isolate the effect of the introduction of blended learning from other key differences between schools that are likely to influence student achievement.

COMPARISON WITH SCORES FROM SIMILAR STUDENTS IN SIMILAR SCHOOLS TAKING THE NWEA MAP ASSESSMENT

This design involved comparing the spring 2012 NWEA MAP scores for students enrolled in the blended learning schools with those of a comparison group of similar students enrolled in similar schools across the nation. NWEA applies an algorithm to create a virtual comparison group (VCG) using scores on MAP adaptive assessments for similar students in similar schools who took the same test during the same test window.[4]

Results of the VCG analysis were used to determine the percentage of students in the blended learning schools that met or exceeded the spring test scores of their VCG and the mean difference in spring test scores between the blended learning schools and the VCG. For the 2011–12 school year, analysis of growth on the NWEA MAP relative to the VCG comparison was limited to KIPP Empower Academy and

[2] *For FirstLine, SRI also attempted to implement a prior-cohort comparison design, comparing students in the 2011–12 grade-level cohorts within Arthur Ashe Charter School with their peers in prior school years. However, because so many prior test scores were missing for the comparison cohorts, this design was not feasible.*

[3] *SRI examined the possibility of implementing this design within the KIPP LA network. However, because of significant differences in student composition between available comparison schools in the two KIPP LA schools available for comparison and KIPP Empower Academy, SRI determined that the design was not feasible.*

[4] *NWEA uses three types of filters to create a VCG for each student being studied: (1) general assessment filters (only students with valid test scores for the same year and subject area), (2) school filters (percentage of students qualified for free or reduced-price lunch and urban/rural classification), and (3) student filters (same grade level, within 5 scale points on the pretest score, and test dates within 14 days of the student). After these three filters are applied, a group of 51 students is randomly selected from the database. The median score of this group is the metric that is used to compare to the study group student score. See http://www.kingsburycenter.org/our-data/virtual-comparison-groups for more information.*

Rocketship schools. Both of these sites implemented the MAP assessment as part of their own formative assessment system.[5]

While this design can provide evidence of the relative performance of the blended learning schools in the analysis, several factors limit the claims that can be made about the direct relationship between blended learning and improved test scores. The primary limitation of the VCG design is that SRI did not have access to VCG comparison data from prior years for the Rocketship schools and KIPP Empower Academy. Thus, it is difficult to know the extent to which a relative difference in spring 2012 test scores is a result of the introduction of blended learning rather than a preexisting difference in performance between the schools implementing blended learning and schools comprising the VCG. Also, little is known about the schools that students in the VCG attended and the instructional programs they experienced. Students in the VCG may have been in schools undergoing instructional reform and integrating technology into their instruction. Therefore, any important existing differences between schools in the study and schools in the VCG that might affect test scores are unknown. What is clear is that the schools served similar compositions of students based on the percentage eligible for the federal free or reduced-price lunch program and administered the MAP assessment during the same test windows.

ACHIEVEMENT GAP ANALYSIS

SRI also examined the extent to which the Rocketship and Alliance schools in the study were able to narrow the achievement gap between low-income and high-income schools in California.[6] We defined high-income schools as other schools in California (K–5 schools for Rocketship and high schools for Alliance) with 10% or fewer of their students qualifying for the federal free or reduced-price lunch program. Low-income schools were those whose proportion of students qualifying for free or reduced-price lunch was within ± 5% of the average of the Rocketship and Alliance schools in the analyses. We compared the difference in achievement gaps (the difference in the mean student performance on the spring 2012 for the two groups of schools) for two sets of school contrasts: (1) all high-income versus all low-income schools in the state and (2) all high-income schools and the schools implementing blended learning included in the analysis. The difference in the gaps for the two contrasts represented the extent to which the blended learning schools were able to affect (reduce or increase) the high-income/low-income gap in test scores that exist in the state. (Further details of the achievement gap analysis are provided in the individual site profiles.)

The primary limitation of the design of the achievement gap analysis is that it could not isolate the effect of the blended learning instruction from the rest of the school's instructional system and other characteristics of the school that may be associated with higher test scores (e.g., leadership, teacher quality, student and family characteristics, and curriculum). Although we have no evidence that the schools in the sample were intentionally enrolling students from low-income communities who were among the most academically prepared and from more supportive families, we expect that families select these schools. To enroll in a charter school, families must seek out and apply to the school, so these families and their students may be different in important ways from families who do not apply to these schools. Because there is no way to reliably control for these differences between students who enroll in charter schools and students who enroll in other low-income public schools, the results of these analyses cannot be used to make definitive claims about the effectiveness of the blended learning models.

[5] *The two other research sites implemented the MAP during the study year—Summit and Alliance—but their MAP scores were not included in any of the analyses reported here. Schools within these sites were implementing the MAP for the first time for the purpose of the research. As a result, teacher buy-in to the value of the assessment for informing instruction varied across the sites and implementation of the fall and spring MAP was uneven, resulting in unreliable growth estimates for participating schools in these sites. In addition, schools within both sites experienced significant technical issues with the implementation of NWEA's online version of the test that disrupted test administration (particularly during the fall/winter test window) and contributed to a less than ideal testing experience for schools, teachers, and students.*

[6] *The achievement gap analyses were limited to ELA and math test scores in Rocketship schools and ELA scores in participating Alliance high schools. In California high schools, the state math assessment a student takes depends on which math course the student is enrolled in during the school year. If students in high-income schools are more likely to take more advanced math courses such as Algebra II or geometry in grades 9 and 10 than students in low-income schools, the majority of whom may be enrolled in Algebra I, then any comparison of Algebra I scores between high- and low-income schools may represent a comparison between the lowest performing students in the high-income schools and some of the better performing students in the low-income schools. As a result, any differences found between the groups may be biased, favoring the low-income schools, and thus difficult to interpret. Math scores thus were excluded from this analysis. Summit schools were also excluded because they implemented blended learning only for math instruction. In addition, because California does not begin state testing until grade 2, KIPP Empower Academy was not included in these analyses .Finally, the analyses were limited to California schools because school-level achievement and free or reduced-price lunch program data for all schools in the state were publicly and readily available through the California Department of Education website. This was not the case in Louisiana, so FirstLine's Arthur Ashe Charter School was not included in the analysis.*

Synthesis of Implementation Findings

Presented here is a cross-site summary of selected implementation findings from the interviews, observations, and teacher surveys at the sites. The findings cover a range of topics including: infrastructure and technology, teacher training and support, elements of the software design, perceived benefits to teaching and learning, and student productivity.[7]

INFRASTRUCTURE AND TECHNOLOGY

Unreliable Internet connectivity, inadequate bandwidth, and problems with software programs hindered many schools' ability to implement their models. Except for the three math teachers at Summit, a majority of the teachers at each site reported that either they or their students experienced technical problems during the year. For teachers at Rocketship, Alliance, and FirstLine's Arthur Ashe, these problems were reported to be either moderate or significant barriers to effective use of their blended learning models. The most commonly cited technical issues were the reliability of Internet connectivity and online programs and insufficient bandwidth to run some of the programs. In one of the most affected sites, FirstLine's director of blended learning estimated that teachers in the Arthur Ashe Charter School lost 3 to 4 weeks of instruction during the school year because of intermittent Internet loss and other technical problems.

On-site IT support and contingent instructional activities provided critical support in the wake of problems with technology. Technology-related problems need to be anticipated and planned for. KIPP LA, FirstLine, and two of the three Alliance schools had hired staff to provide on-site IT support, which freed teachers from this burden. FirstLine's Arthur Ashe had a TeachUp! intern who provided day-to-day on-site IT support and another IT person to take care of larger network issues. At the KIPP Empower Academy, a technology specialist roamed between classrooms throughout the day to ensure that all hardware and software programs were working properly and to assist students who were having difficulty logging in or other technical problems. Even with on-site IT staff available, administrators, teachers, and lab monitors in several sites expressed the need for backup activities and software programs for the times technology issues arose. Examples of such activities in lab rotation models were the use of worksheets or off-line projects or the use of software programs that resided on the students' computers or schools' servers.

Site-based blended learning coordinators and teams provided critical technical support for administrators and teachers during the transition to blended learning. These positions were essential to reduce some of the burdens on administrators and teachers so that they could stay focused on teaching and learning. People in these coordination roles were often responsible for identifying and piloting online programs, negotiating with vendors, providing reports on student performance in the online programs, coordinating professional development, and overseeing the purchase of technology. In some cases, such as in FirstLine and KIPP LA, the coordinator served at the network level and supported all schools in the network; in others, such as Alliance's Tennebaum Family Technology High School, the person served as the coordinator for that school specifically. Within Rocketship, curriculum and technology specialists operating out of the national office were responsible for the design of the learning lab and its curriculum as well as the selection and ongoing evaluation of the online programs used in the lab. In Rocketship schools, the assistant principals were responsible for establishing the learning lab culture and supervising the learning lab staff. The role of the project manager at KIPP LA was to coordinate the implementation of the blended learning model at KIPP Empower Academy, including negotiating with external vendors and consultants and tracking progress against project timelines.

TEACHER TRAINING AND SUPPORT

Teachers' satisfaction with training associated with the adoption of the blended learning model varied by site. Teachers at all the schools reported participating in a training or orientation directly related to the school's use of blended learning or on the software programs supporting it. Approximately two-thirds of teachers at Alliance and Rocketship and one-third of those at FirstLine's Arthur Ashe Charter School, however, indicated they were dissatisfied with the training. The reasons they cited included the format (via phone or the Internet), which lessened the effectiveness of the training, and training materials that were too theoretical and not specific to particular content areas. Teachers who reported being satisfied with the training noted that the training was in person, led by other teachers rather than administrators or vendor representatives, had hands-on activities, and, when the training was on specific software, that a company representative was accessible and available to answer

[7] *Due the variation in the blended models adopted by each research site, findings related to the impact of the various blended learning models on student learning and the schools' ability to close the high-income/low-income achievement gap are reported in each site's research profile.*

questions during the school year. Several teachers interviewed said they would have benefited from additional training during the school year, indicating that the initial training provided a good introduction but that training would have been more beneficial once they had some familiarity and experience with their blended models.

The need for teacher training varied by the type of blended learning model adopted. The roles teachers play in the implementation of the online instruction component of the blended learning model varied greatly across the models, with each having different expectations for teachers' interactions with the software and use of system data on student progress. Therefore, the level and nature of training teachers received on the operation of the software programs also varied. Teachers involved in the use of the online instruction component of the model—those at Alliance, Summit, and FirstLine—may have required more training than teachers who were not expected to interact with the online programs at sites like Rocketship and the KIPP Empower Academy. Yet even teachers in Rocketship schools reported on the survey that they believed it was *very important* to receive training on how to access and interpret student progress reports provided by the online instruction programs (90%) as well as how to use the data to inform their instruction (75%).

SOFTWARE DESIGN ELEMENTS

During 2011-12 school year, sites continued to experiment with online instructional software in ELA and math to best serve their students and teachers. They continually refined their models, which included searching for and piloting different online programs. In general, sites were interested in identifying programs that (1) were comprehensive and aligned with the Common Core standards, (2) were adaptive, (3) were interactive and engaging, (4) were assignable, (5) provided valid reports of student learning, (6) allowed for integration into the school learning management systems and had a single sign-on portal when more than one program was used, (7) were cloud based, and (8) were affordable.

How assignable online instruction programs are may limit their integration with the classroom curriculum. Many sites used online instructional programs such as DreamBox Learning and Istation that are adaptive and enable students to work at their own pace on content that is appropriately challenging. Given this self-pacing, students may not be working on the same topic at the same time or at the same time the teacher is presenting the material in the classroom. Many of the programs do not allow teachers or administrators to assign the online content for students. During the 2011–12 school year, Rocketship attempted to address this by working with its program vendors to sequence the content delivery so that it was roughly aligned with the schools' curriculum pacing guide. This required mapping the program's content against the content standards and sequence the schools follow, something that most teachers reported is frequently lacking. The importance of program assignability to the implementation of a blended learning model depends on the role of online instruction in the model. That is, assignability is less critical if the goal for the use of the online programs is remediation, filling in students' knowledge gaps and catching them up to their peers, as in the KIPP Empower Academy classroom rotation model. Assignability is more critical when the online instruction is to be an integrated supplemental activity directly supporting the lessons the teacher delivers in the classroom.

Teachers reported that the quality of the online software programs varied by subject. Across sites, teachers and administrators tended to be more satisfied with the content coverage and quality of software programs developed for math skills than for ELA. For ELA, the concerns were greatest about supporting writing instruction. For example, Alliance leadership expressed concerns over the inadequacy of short written-response assignments, prevalent in many programs, to prepare students for college-level writing.

Several factors limited teachers' use of data from online systems to inform instruction. All the sites participating in the research had strong cultures of using student performance data to make informed instructional decisions. A majority of teachers surveyed reported reviewing student performance data from the online programs, but their reports of both frequency and the value of the data varied. Approximately half reviewed student reports at least once per week, while the others reported reviewing student data no more than three times per month. Some schools have time set aside on the schedule for regular reviews of the data from the online programs; for example, FirstLine's Arthur Ashe Charter School had Data Days every Wednesday that gave teachers a built-in opportunity to review system data, identify students who are struggling, and determine

appropriate next steps or interventions. Within Rocketship, the lab monitors (or individualized learning specialists) and school leaders were responsible for reviewing the data; they held weekly meetings to review student progress relative to weekly progress goals and met bimonthly to identify a focal group of students who needed additional support.

In general, however, schools and teachers have underutilized data from the online instructional programs. Teachers and administrators noted several challenges in leveraging the student data. One challenge related to the quality of dashboards used to report student progress in some of the programs. Teachers reported that either the dashboards were difficult to access or the information on student progress was difficult to interpret in terms of the schools' own content standards, which in turn hindered teachers or lab monitors in efficiently identifying students who were struggling or falling behind progress expectations. Teachers also reported that progress and performance measures were often not aligned with state content standards or the schools' curricula, leaving them less able to anticipate how students might perform on state-mandated assessments or internal benchmark exams.

Teachers' and administrators' trust in the assessments embedded in the online programs was also limited; several reported they trusted their own formative assessments more than a program's judgment of content mastery because they questioned the validity of the program's measures. Those individuals were unclear about how well content "mastery" as defined by the program aligned with their own measures and notions of mastery and those associated with state-mandated assessments.

Several sites that used more than one online program in their blended learning model faced the challenge of integrating multiple streams of student performance data from the programs into their own student information systems. Because the data from the online programs were not well integrated into a single dashboard, many teachers found that reviewing the data was too time consuming to do on a regular basis. Third-party providers such as Education Elements have entered the market to support schools with this. (Education Elements worked with KIPP LA and Alliance during the 2011–12 school year.) However, during the study, the research sites were left to make sense of the multistream data on their own. For example, the blended learning coordinator at FirstLine estimated he spent 2 hours per week aggregating student progress across the different online programs Arthur Ashe Charter School used (each of which reports progress in different ways) into a single report for school staff. Since each of the programs reported student progress in different ways, the coordinator developed “report interpretation guides” so that school staff could efficiently interpret the meaning of the data. Without the coordinator’s efforts, the usefulness of the system data for monitoring students’ progress in the learning lab and identifying students in need would have been severely limited.

BENEFITS TO TEACHING AND LEARNING

Sites are using blended learning models to personalize learning through self-paced programs, adaptive online instructional content, and facilitation of small-group instruction for students with the greatest academic needs. Administrators and teachers described several uses and benefits of blended learning to personalize students’ learning. Self-paced and adaptive math software programs like DreamBox Learning, ST Math, and Istation support personalized learning by individualizing instruction, challenging students with the appropriate level of content, and redirecting their path through the program based on prior successes in the program. Students progress through the material at different speeds based on their learning needs. Students may take more time to complete a given topic, skip topics they already know, repeat topics they need help on, and explore content that is beyond the curriculum. A majority of the teachers surveyed reported that technology and computer-mediated instruction played a major role in providing enrichment for advanced students and remediation for struggling students.

Learning is further personalized within all the sites by using blended learning to facilitate small-group instruction. For example, at FirstLine’s Arthur Ashe Charter School, administrators reported that their learning lab model gives teachers flex time to include more one-on-one and small-group instruction. Under this model, subject-specialty core teachers and special education teachers could remain in the learning lab to provide students with moderate academic needs (also known as Tier 2 students) in the school’s response-to-intervention (RtI) program with additional instruction, either in the lab (push in) or in a nearby classroom (pull out). During this time, these students worked on different online programs based on their needs. For example, students who needed more remediation were assigned to work on ST Math while students who were ready for more challenging material used Think Through Math. Students with the greatest academic needs at Arthur Ashe—RtI Tier 3 students and special education students—spent their lab time in a “learning support” classroom, rotating in small groups between teacher-led activities, online workstations, and independent practice. Similarly, at Rocketship the learning lab was key in supporting students with the greatest academic needs. Every 8 weeks, benchmark assessment data were used to identify students to receive small-group tutoring from lab monitors during the bulk of the learning lab time while others used the online programs.

KIPP Empower Academy’s in-class rotation model also enabled small-group instruction in the ELA classroom. Students were grouped by ability, and teachers differentiated their instruction based on the level of each group. As one group of students worked on adaptive online software programs, other students received instruction from teachers in the classroom. Students typically received instruction, online or teacher led, in groups of 14 or fewer. Similarly, in Alliance’s in-class rotation model, teachers provided direct instruction to small groups (typically 16 students or fewer) at one of several learning stations, including an online instruction station.

Teachers reported that blended learning benefits students’ procedural skills development more than higher order thinking. Perhaps because of online instructional programs’ emphasis on providing students with opportunities to practice basic skills and procedures, a majority of teachers at most sites reported that their blended models were more effective at helping students recall basic facts than at helping them develop higher order thinking skills. This finding held true across subject areas. Teachers’ perceptions of skill development in various areas reflected their schools’ blended learning model as a whole—online and offline instruction—and not just the online or computer-mediated component.

Goal-setting promotes self-directed learning in the blended learning environment. The goal-setting practices associated with many of the blended models promote students’ sense of ownership over their learning and accountability. Most often, goal setting was associated with students’ weekly goals for progress on the online programs. This was the case for FirstLine, Summit, and Rocketship. Goals were set and communicated by the lab monitors in FirstLine and Rocketship and by the teachers in Summit. In Alliance schools, teacher-developed weekly "digital agendas" specified what online and offline tasks a student needed to complete each week. A majority of administrators, teachers, and lab monitors interviewed reported that weekly goal-setting helped students to become more invested in their learning and to see the rewards and consequences of not meeting goals. For example, at Arthur Ashe Charter School, students who fell far behind weekly progress expectations were required to miss a school wide Friday activity and instead spent the time completing their work in the learning lab. Similarly, Summit required students to complete their weekly Khan Academy problem sets or remain after school on Fridays to do so; completion of the problem sets contributed to students' math grades.

Teachers surveyed reported that students' readiness for self-directed learning may vary by their academic preparation. While teachers cited advantages with their school's blended learning model, they generally indicated that students did not reap these benefits equally. Across sites, a large majority of teachers reported that their models were at least *somewhat effective* for all students but that those who benefited most were students whose academic work was either ahead of, or at the same level as, most students their age; the models were possibly less effective for students whose academic work was below their expected grade level. In addition, many teachers interviewed felt it was the ability to self-manage and self-direct their learning that determined which students would most likely thrive in a blended learning model.

Teachers' satisfaction with their school's blended learning models varied. Most teachers at FirstLine, KIPP Empower Academy, and Summit reported they were satisfied with blended learning in their instruction and would recommend it to other teachers. However, approximately half the teachers at Alliance and Rocketship reported they were less satisfied. Data collected during site visits and responses to the teacher survey suggest some explanations for these differences. Within Rocketship schools, teachers by design were largely disconnected from what took place within the learning lab, and this may have influenced their feelings about the online instructional component of the model. A Rocketship teacher interviewed best exemplified this sentiment, commenting that although she knew her students were working hard in the learning lab, she did not know anything about what they were learning or how to connect what they learning to her instruction in the classroom. Many Alliance teachers spoke of the classroom management challenges in implementing an in-class rotation model with large class sizes and high school students. Other teachers at Alliance, particularly those in subjects other than math, also reported that the three-station model (direct instruction, independent, and collaboration) was perhaps not ideal for the content they were teaching. Several of these teachers expressed frustration at having to drop lesson plans they had found to be successful in the past to conform to the station model.

STUDENT PRODUCTIVITY

Strong classroom/learning lab management practices are critical to ensure student productivity in online environments. Teachers and lab monitors alike stressed the importance of establishing the proper academic culture, norms, and behavior management practices for a blended learning model to be successful. This is especially important in managing student transitions between online and teacher-led instructional blocks, particularly with younger children. At FirstLine's Arthur Ashe Charter School, students were expected to move to their seats in the learning lab quietly and efficiently, store their bags neatly under their desks, log in to the software promptly, and sit up straight in their chairs with headphones on (as appropriate). In classroom rotation models, similar expectations were established within some sites to promote seamless transitions between stations. For example, at KIPP Empower Academy, during the second week of the school's regular summer school session, students repeatedly practiced the ritual of moving from one station to the next following a typical rotation schedule. Students practiced timed transitions with a goal of moving from one station to the next in under 60 seconds to maximize instructional time. During the first 2 weeks of the school year, teachers also spent extra time modeling the behaviors students were expected to follow when they were in the online instruction station, including how to sit, where to place their hands, and how to care for the headsets, as well as the procedures for logging in and out.

ESTIMATING THE IMPACTS OF EARLY ADOPTION BLENDED LEARNING MODELS ON LEARNING

Interpreting evidence of impacts with caution. The evidence of learning impacts presented in the individual site profiles that follow are preliminary and cannot be used to make definitive claims about the effectiveness of any of the blended learning models implemented during the time of the study. For all the sites, except Rocketship schools, the 2011–12 school year was their first or second year of experimenting with blended learning instruction. Even in Rocketship, for two of its five schools the 2011–12 school year was the first year of

Michael & Susan Dell FOUNDATION

TABLE 3: QUASI-EXPERIMENTAL DESIGNS USED TO EXAMINE IMPACT OF BLENDED LEARNING MODELS ON STUDENT'S LEARNING

RESEARCH SITE	WITHIN-NETWORK COMPARISON	VIRTUAL COMPARISON GROUP
Rocketship	No	Yes
Summit	Yes (math only)	No
Alliance	Yes	No
KIPP Empower Academy	No	Yes
FirstLine	Yes	No

operation and for another school it was the second. Each of the sites refined its model in different ways throughout the year to best meet the needs of teachers and students, including adjusting the time allocated to different online programs and, in some cases, stopping the use of or replacing online programs that were not meeting expectations. By the end of the 2011–12 school year, many of the sites were planning further alterations to their models for the 2012–13 school year. In addition, teachers' use of student learning data collected by the online instructional programs to inform instruction was still in its infancy. According to site administrators, at the time of the study solutions were not in place that provided teachers with data that was both easy to interpret and actionable. Thus, impact estimates based on the models implemented during the 2011–12 school year must be interpreted in the context of their specific and early adoption experiences and may not reflect the true potential of blended learning in these sites.

Last, as described above, the impact estimates are based on analytical designs that had limitations. Primarily, even though the designs applied were the most rigorous available to the researchers at the time, they could not isolate the impacts of the blended learning models from other aspects of the schools that might also affect learning, such as differences in curriculum, teacher quality, and the academic culture. The quasi-experimental designs SRI used to collect evidence on the impacts of blended learning use in each site are summarized in Table 3.

Interpreting differences in test scores. To aid in interpreting the differences in test scores across sites, tests, grade levels, and subject areas, we report the difference in adjusted mean scores as a standardized effect size. An effect size expresses the difference between two scores in terms of how spread out the scores are. (Technically, the effect size is expressed in terms of standard deviations of outcome scores).[8] An effect size of 0.3, for example, means that one group on average scored 0.3 standard deviations more than the comparison group. This would apply whether the scale of the test score were 0 to 100, 150 to 600, or any other measure. That is, an effect size of 0.3 would essentially represent the same magnitude of difference regardless of the underlying point system used by the outcome measure. Because of this property, researchers commonly use effect sizes to compare the impacts of interventions across different tests. In addition to reporting effect size for each outcome measure analyzed, we translate the effect sizes into the expected percentile gain for the median student in the comparison group (the student who scored at the 50th percentile on the outcome measure) if that student had attended one of the schools where the blended learning models had been implemented. This is also known as the improvement index. For example, an estimated impact with an effect size of +0.25 standard deviations means that the median student in the comparison group would have scored at the 60th percentile, or an improvement of 10 percentile points, if that student had attended one of the schools in the research site that had implemented the blended learning model.[9]

[8] *An effect size is commonly computed by taking the mean difference in test scores between the treatment and comparison groups and dividing that difference by the pooled standard deviation for the total sample (treatment and comparison students combined).*

[9] *For additional details on the computation of the improvement index see What Works Clearinghouse Procedures and Standards Handbook Version 2.0, December 2008 (http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_standards_handbook.pdf)*

Implications and Future Research and Development

Given the schools' early-stage adoption of blended learning and the continuing evolution of the models at each site, it is much too early to judge the effectiveness of these models relative to other ways to organize instruction. The preliminary evidence collected during this research, however, is encouraging. Students attending a majority of the schools in the study where comparative designs could be applied outperformed students in the comparison schools not using blended learning. In those schools, blended learning is clearly an important component of what appears to be an effective instructional system. However, as described, the designs used to collect this evidence have limitations that prevent us from attributing differences in learning outcomes to blended learning alone. As new and refined blended learning models emerge and show promise, investments should be made in implementing rigorous evaluation designs to carefully study their benefits at scale, determine how the benefits vary by types of students and why (e.g., motivation, academic preparation), and identify the supports and design features needed to help all teachers and students succeed in a blended learning environment.

During this research, several important facilitating factors emerged that will most likely deserve consideration by others looking to adopt blended learning models like those implemented by schools in the study.

> *"Setting weekly goals for student progress as a mechanism to focus them and increase productivity was a practice that emerged across the school year in most of the schools in the study."*

Blended learning coordinators played an important role in supporting schools' adoption of blended learning. Each of the research sites had someone or a team of people who supported the schools in their adoption of blended learning, freeing the administrators and teachers of the logistical burdens. This person or team was responsible for identifying promising online programs; arranging for schools to pilot the software and teacher training; negotiating with software and hardware vendors, consultants, and Internet providers; and in some cases, generating teacher-friendly reports of students' progress on the online programs. Without such support, these schools probably would not have had the success they did in this early adoption cycle.

Establishing productive self-directed learning cultures is important for students to fully benefit from online learning. Students as young as 5 and 6 years old are spending up to a few hours each day in online instruction in a number of schools in the study, those associated with Rocketship, FirstLine's Arthur Ashe Charter School, and KIPP Empower Academy. Each of these sites had spent considerable time over the prior summer and throughout the school year helping students make efficient transitions between regular classroom instruction and their online instruction time and establishing behavioral norms for the learning lab or online station. Staff assigned to manage the learning labs also spent considerable time each day keeping students focused on the task at hand and reminding them about the proper behaviors while working on the online programs. Staff in many of these schools also designed activities to promote better productivity in the learning labs by creating friendly competitions among students and classrooms and by recognizing students and classrooms that had achieved a staff-set weekly goal for online progress or the highest point total on a program. Setting weekly goals for student progress as a mechanism to focus them and increase productivity was a practice that emerged across the school year in most of the schools in the study.

Single sign-on portals allow even very young children to quickly access online programs. For in-class and lab rotation blended learning models, efficient transitions of students between the teacher-led instruction and online instruction are critical for productive use of instructional time. Reliable single-sign on portals were an important feature of efficient transitions. For example, children in kindergarten and first grade attending the KIPP Empower Academy were able to transition to the online station and begin working within a program such as DreamBox Learning within 30 seconds after completing an activity in a teacher-led station. This was possible not only because these students spent a part of their summer school camp practicing the transition, but also because the single-sign on portal Education Elements developed allowed for easy access. The portal enabled students to quickly log in and access the entire portfolio of online programs using a single graphical username and password. Rocketship and FirstLine also used single-sign on portals to support efficient transitions between the classroom and their learning labs.

Adequate on-site IT support, reliable Internet connection, and sufficient bandwidth, are critical technology-related factors for effective implementation. Schools implementing blended learning must anticipate and plan for technology-related problems that can interfere with the delivery of instruction. Problems that arose during the school year included significant reductions in bandwidth during login and periods of high video use, loss of Internet connectivity, problems associated with the reliability of single-sign on portals, and intermittent issues with accessing cloud-based programs. Across all schools, teachers reported that technology issues greatly reduced the schools' ability to effectively implement their models. All sites had on-site IT support available to support teachers, but few schools had alternative instructional plans ready for when access to online programs was not available, such as providing students with access to software or apps on local computers or servers when Internet access was lost.

A majority of teachers and staff managing learning labs reported that insufficient bandwidth and unreliable Internet connections were significant barriers to their effective adoption of blended learning. The poor reliability of high-speed connections will become an even greater impediment as online instruction becomes more video-rich and bandwidth dependent. Learning lab-based models are particularly susceptible to this problem because from 50 to 150 students may be accessing the Internet simultaneously. Under these conditions, a loss of the Internet or insufficient bandwidth can be very disruptive to schools' instructional plans and student learning.

FURTHER RESEARCH AND DEVELOPMENT

As more schools plan to adopt blended learning models, research and development efforts should consider the supports teachers and students need to more effectively benefit from blended learning environments.

Online programs must be designed to support integration of the content into the curriculum. Some developers such as Khan Academy and DreamBox Learning are actively partnering with districts, CMOs, and independent schools as they seek to refine their programs in response to feedback from administrators, teachers, and students. For schools and teachers to make better use of it in their instruction, online content needs to be designed and organized so that it can be flexibly integrated with regular teacher-led instruction. At a minimum, the content needs to be aligned with the school's curriculum standards or allow teachers or curriculum coordinators to easily map the content to their local standards. The content thus must be searchable and assignable and clearly mapped to grade-level state content standards.

Reports of student progress within the online programs must be easy for administrators and teachers to access and interpret. Developers and system integrators need to address several challenges to realize the promise of online learning programs as a tool to collect and report real-time learning data to students and teachers. First, administrators and teachers must be able to trust that the online programs' measures of content mastery (when available) are valid and can be used to inform instruction and identify students who are struggling and need more support. For the majority of the systems in this study, this was not the case. The sites were using student data from the systems primarily to monitor students' progress through the online instruction as an indicator of their engagement and productivity. The reports were not being used as a source of information about what students were learning. To monitor student learning, most administrators and teachers continued to rely on their own formative assessments.

Underutilization of the online data is compounded when multiple online programs are being used (a common approach for sites in the study), each with its own way of measuring and communicating student performance and progress. The burden for administrators and teachers in interpreting and making sense of these multiple data streams for instructional purposes is too great, so potentially valuable real-time learning data are not being used. The hope is that interoperability standards will one day make it easier for schools to integrate such data with their learning management systems; but until there is standardization of what constitutes mastery across online programs providing instruction in the same subject area, the potential benefits of having ready access to real-time learning data will most likely go untapped. One approach to standardization may be through developers' universal adoption of evidenced-centered-design (ECD) practices for development of the embedded assessments

Michael & Susan Dell FOUNDATION

within online programs (Mislevy, Almond, & Lukas, 2003). The use of ECD practices would ensure that developers are using a common and principled design architecture to develop measures of the same underlying knowledge and skills of interest.

Finally, we need a better understanding of how to help all students succeed in self-directed online learning environments. Some students will probably be better prepared to succeed in online learning environments than others. It is also likely that a single online program or format will not meet the needs of all students, just as the effectiveness of different face-to-face instructional formats (e.g., teacher lectures, peer collaboration, use of worksheets) is likely to vary across students. In addition, in some underserved communities, students are entering schools with relatively low digital literacy skills and are being asked to spend a significant part of their day in a computer-mediated learning environment. As more and more students are asked to self-direct their learning, even children as young as kindergarten-age, little in their prior schooling or home experiences may have prepared them for this. A large percentage of the students currently being exposed to self-directed digital learning environments will need to be taught what it means to manage their own learning and how to seek help from resources other than the teacher, such as online resources. This is particularly important given the movement toward the use of more online programs (including MOOCs) in secondary and higher education and the notoriously low completion rates associated with them. Research is needed to understand the role of noncognitive student characteristics—e.g., motivation, persistence, resourcefulness—in students' success in self-directed online learning environments and the features of online learning programs that better support learners with different levels of preparation. More emphasis is needed within the research to understand who is flourishing and failing in these environments and why and then develop programs and blended learning models with the supports that help all students succeed.

References

Means, B., Toyama,Y., Murphy. R., Bakia, M., and Jones, K. (2009). Evaluation of evidence based practices in online learning: A meta-analysis and review of online-learning studies. Washington, D.C.: U.S. Department of Education.

Mislevy, R., Almond, R., & Lukas, J. (2003). A brief introduction to evidence centered assessment design. Educational Testing Service, Princeton New Jersey.

Staker, H. & Horn, M. (2012). Classifying K-12 blended learning. Clayton Christensen Institute for Disruptive Innovation: San Mateo, CA.

Michael & Susan Dell FOUNDATION

Alliance College-Ready Public Schools Profile (2011–12 School Year)

Introduction

Alliance College-Ready Public Schools (Alliance), founded in 2003, is a charter management organization (CMO) that currently includes 15 high schools and 7 middle schools distributed across roughly 100 square miles of the Greater Los Angeles region. Alliance's mission is to open and operate a network of small public middle and high schools in historically underachieving, low-income communities and have these schools consistently demonstrate strong student academic growth and graduate students ready for success in college.

Alliance began implementing the Blended Learning for Alliance School Transformation (BLAST) instructional model at two high schools—Cindy & Bill Simon Technology Academy High School (Simon) and Judy Ivie Burton Technology Academy High School (Burton) —during the 2010–11 school year, and expanded to a third high school—Tennenbaum Family Technology Academy High School (Tennenbaum)—during the 2011–12 school year (the school's first year of operation). The primary focus of the BLAST model, which is described in detail in the next section, is to support student-centered learning through small-group, data-informed instruction.

BLAST MODEL DESCRIPTION

During the 2011–12 school year, Alliance implemented a classroom station model for core subjects (English Language Arts [ELA], math, science, and social studies). Each classroom contained up to 48 students typically divided into three groups of 16 leveled by ability. Core subject areas met for 120 minutes, 2 days per week, and 47 minutes on Wednesday (a half-day). Each group of students typically rotated through each of three 40-minute stations during the 120-minute periods: teacher-guided instruction, adaptive and online individual instruction using laptops and learning management systems, and project-based collaboration and discussion. There was some flexibility and teacher discretion over the number of stations, however, which ranged from two to four depending on the nature of the day's activity and/or the subject area. At their discretion, teachers also incorporated whole-group direct instruction or activities, most often at the beginning or end of class. Each group had a teacher-designated student leader/facilitator who was responsible for making sure groups stayed on task and smoothly transitioned between stations.[10]

In each core course, a downloadable "digital agenda" was developed each week by the teacher to guide individual student work during the online station. The digital agenda included the learning goals, online tasks, and daily assessments for the week that needed to be completed. Each task included a brief description of what the teacher wanted the student to accomplish in the task along with the link to the digital resources. Each student within a leveled group received the same learning goals and tasks to complete each week along with individual-specific tasks tailored to each student's learning needs.

One school, Tennenbaum, also implemented an online learning lab model in which students enrolled in online courses in noncore subject areas (e.g., Spanish, Psychology) and in core areas to recover credits. Students were in the learning lab for 120 minutes, 2 days per week and on Wednesdays for one 47-minute period. Each learning lab contained up to 96 students taking different courses delivered using various modalities, including distance learning and online courses, via desktop computers. During the lab time about 65% of students were enrolled in either an online California High School Exit Exam (CAHSEE) preparation class provided by Revolution K12 or an online credit recovery course, primarily in ELA or math. Other students used their time in the learning lab to participate in online courses taken as electives.

Extra instructional sessions were also offered after school and on Saturday. The after-school sessions met 4 days per week for 2 hours. A Saturday Academy was held three weekends per month, for four hours and focused on ELA and math instruction. While some students were selected to attend after-school sessions based on their needs, all students were required to attend Saturday Academy (attendance rates were about 75% according to Alliance leaders). During these extra support sessions, a variety of instructional modalities were provided (one-on-one tutoring, small group teacher-led instruction, and online instruction). The modality used during the after-school session was determined by the teacher based on a student's needs. During the Saturday Academy, students got to select the instructional modality and worked in this mode for the duration of the session.

Special education students also received additional supports both in the classroom and during the learning lab time. Instructional specialists "pushed in" to the classroom and provided support to special needs students during the various rotations in the classroom. Within the online learning lab period, some students, depending on their needs, were "pulled out" into a resource lab and received additional one-on-one time with the specialists.

Alliance used a combination of online learning instructional programs to support the implementation of their blended model. Compass Learning and Revolution K12 are the primary providers for online math instruction. Other online resources to support math included Khan Academy and Brightstorm. Revolution K12, an adaptive, self-paced program, was used to "backfill" gaps in students' foundational skills and to prepare 10th graders for the CAHSEE in math and ELA. For online ELA instruction, primary providers included Compass Learning, Revolution K12, and Achieve3000. Achieve3000 provided targeted leveled reading instruction. Some students, particularly within Tennenbaum, also had access to Apex Learning courses for credit recovery and to fulfill elective requirements.

Students in the three schools had access to one-to-one computing. At Tennenbaum and Simon, each student checked out a laptop at the start of the school day, carried it with them from class to class, and returned it at the end of the day. At Burton, laptops were stored on carts within each classroom; students selected a laptop to use for the duration of that class and returned it at the end of the period. Tennenbaum also had a computer lab with desktop computers where students engaged with Apex courses or prepared for state exams.

[10] *A detailed description of Alliance's instructional and operational model is available in a case study developed as part of this research study and published on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).*

Sample

Table 1 summarizes the study sample for the three Alliance College-Ready schools implementing the BLAST instructional model (Tennenbaum, Simon, Burton) and the three Alliance schools in the comparison group that were not implementing the BLAST model (William and Carol Ouchi High School, College-Ready Academy #5, Health Services Academy). For school year 2011–12, Alliance had 14 charter high schools in its network, all located in the Los Angeles Unified School District. Simon and Burton first adopted a version of the BLAST model during the 2010–11 school year, while Tennenbaum's first year of implementing a BLAST model was the 2011–12 school year. The three Alliance schools used as the comparison schools did not adopt the BLAST model during these years.

Similar school rankings (when available) range between 3 and 10 (with 10 being highest) and indicate the schools' baseline performance based on spring 2011 state test scores (Tennenbaum opened in fall 2011 and thus did not have a ranking). Simon is the only school that performed below the average high school in its grouping. The other schools—the two other BLAST schools and the comparison schools—were all performing near the top of the state relative to their peers. Academic Performance Index (API) based on spring California Standards Test (CST) scores for the 2011–12 school year for the BLAST schools ranged from 628 to 737 with an average of 668, while API scores for the comparison schools ranged from 713 to 773 with an average of 750. An API score of 800 is the statewide goal set by the California Department of Education.

TABLE 1. ALLIANCE BLAST SCHOOLS AND COMPARISON SCHOOLS IN SAMPLE

SCHOOL NAME (YEAR OPENED)	GRADE LEVELS	NUMBER OF STUDENTS	2011–12 API SCORE[A]	2010–11 SIMILAR SCHOOL RANKING[B]
BLAST SCHOOLS				
Tennenbaum Family Technology Academy (2011)	9–11	450	628	NA
Simon Technology Academy (2010)	9–11	450	639	3
Burton Technology Academy (2005)	9–12	600	737	10
NON-BLAST COMPARISON SCHOOLS				
Ouchi (2006)	9–12	550	773	10
College Ready Academy #5 (2007)	9–12	605	765	10
Health Services Academy (2009)	9–12	500	713	9

[A] *The API ranges between 200 and 1,000, with 800 as the statewide goal for all schools.*

[B] *To create the Similar Schools Ranking the California Department of Education compares a school's test scores to 100 schools across the state with similar demographic profiles. California uses parent education level, poverty level, student mobility, student ethnicity and other data to identify similar schools. Rank 1 means the school performed below at least 90 of its 100 similar schools. Rank 10 means the school performed above at least 90 of its 100 similar schools.*

TABLE 2. DATA COLLECTION SUMMARY, FALL 2011 AND SPRING 2012 SITE VISITS	FALL 2011 SITE VISIT	SPRING 2012 SITE VISIT
OBSERVATIONS		
Schools	1[A]	3[B]
Classrooms	3	11
INTERVIEWS		
CMO Leaders	3	2
School Leaders	2	3
Teachers	4	14
Students	3	12

TABLE 3. RESPONSE RATES, TEACHER SURVEY, ALLIANCE

INSTRUMENT	RESPONSE RATE
BLAST Teacher Survey	85% (22 of 26)
Comparison Teacher Survey	66% (23 of 35)

Data Collection

SITE VISITS

One member of the research team conducted a site visit to Alliance in October 2011 and two members of the team visited again in May 2012. The fall site visit was conducted over 2 days and the spring visit over 3 days. Table 2 summarizes the number of classroom observations and interviews conducted during each site visit. For the fall 2011 site visit, the SRI team only visited Tennenbaum, but for the spring 2012 visit the team visited all three Alliance schools implementing the BLAST model.

TEACHER SURVEY

SRI conducted surveys of teachers implementing blended learning in the BLAST schools as well as teachers in the comparison schools. Table 3 shows the response rate for each survey, with 85% of BLAST teachers responding (22 of 26) and 66% of comparison teachers responding (23 of 35).

SYSTEM LOG DATA

In addition to the teacher survey data, system log data was provided by Alliance for students in the three schools implementing the BLAST model in SY 2011–12. For the purpose of the system use analyses, only students' interactions with the online instructional programs within ELA and math classrooms were analyzed; we did not collect use data for science, history, or foreign language due to variability in teachers' use of online resources and materials to support these subjects. Therefore, our analysis of students' use of math and ELA programs does not represent the full extent of technology-supported instruction a student received during the school day. We also did not capture students' use of online resources outside these programs, including tools to conduct online research, collaboration tools, and programs used to produce learning products (word processing, multimedia production) and analyze data. Nor did we have access to data from students' use of free open education resources like Khan Academy. Finally, we also did not analyze the amount of time students spent in Apex Learning courses. Less than 15% of students in Tennenbaum and Simon enrolled in an Apex Learning ELA or math course and there were no student enrollments for Burton. Table 4 shows the core online programs that were included in the use analysis for ELA and math.

TABLE 4. CORE ONLINE PROGRAMS USED FOR ELA AND MATH INSTRUCTION

ELA	MATH
Compass Learning	Compass Learning
Revolution K12	Revolution K12
Achieve3000	

[A] *Tennenbaum.*

[B] *Tennenbaum, Simon, Burton.*

Michael & Susan Dell FOUNDATION

TABLE 5. OUTCOMES MEASURES FOR IMPACT ANALYSES, ALLIANCE

	ELA	MATH
Grade 9	2012 ELA Benchmark 2012 CST ELA	2012 CST Alegebra I[A]
Grade 10	2012 ELA Benchmark[A]	NA

For the three online programs used to support ELA and two for math instruction, we analyzed the amount of time that students logged on each system. This was the only measure that was common across the three systems. Since Compass Learning and Revolution K12 include content in math, ELA content, and other content areas, the content students worked on was coded as either math, ELA, or other, and corresponding time on math and ELA content was computed. For Compass Learning, courses were made up of multiple activities, and minutes were reported for each student by activity type. The total time students spent working on Compass Learning was computed by aggregating time across activities. For Revolution K12, the total number of hours that a student spent in a course was reported. Minutes for Achieve3000 were reported on a monthly basis associated with the amount of time that a student spent on assessment activities with Achieve3000; these monthly values were summed for each student.

STUDENT LEARNING OUTCOMES

Table 5 lists the different achievement outcome measures used in the analyses of impacts on student learning for Grades 9 and 10. Measures of learning outcomes included the California Standards Test (CST) for ELA and math (Algebra I) and Alliance's own internal ELA benchmark assessments (scale of 1–5).[11]

Implementation Findings

Findings from the site visits and teacher survey were used to help understand the facilitating factors and barriers to implementing the BLAST model during the 2011–12 school year. Additionally, the student log data from the online programs was used to better understand the intensity with which students used the primary online programs supporting the BLAST model for math and ELA instruction.

FINDINGS FROM THE SITE VISITS AND TEACHER SURVEY

Data from the fall 2011 and spring 2012 site visits and the teacher survey were used to report on different aspects of implementation of the BLAST model, including infrastructure and technology issues, the support and training teachers received, characteristics and quality of system software, and teachers' use of the system data to support their instruction. In addition, we report on teachers' overall satisfaction with the BLAST model and the perceived benefits to students and teaching of the BLAST model as reported by teachers and through the research team's own observations. A summary of the survey results is included in the Appendix for this profile.

[11] *We also intended to analyze 10th-grade pass rates on the California High School Exit Exam. However, due to a high level of missing prior achievement data for 10th-grade students, this outcome measure could not be included in the analyses. In addition, we also intended to analyze the Northwest Evaluation Association Measure of Academic Progress for math and ELA assessments. Technical difficulties associated with the online administration of the tests resulted in an uneven implementation of the assessments across schools and thus rendered the results of the assessments invalid for analysis purposes.*

[A] *Sample does not include students in the Tennenbaum Family Academy due to the high levels of missing data (greater than 25% of the sample) for prior CST scores.*

Difference in instruction between BLAST and Non-BLAST Schools

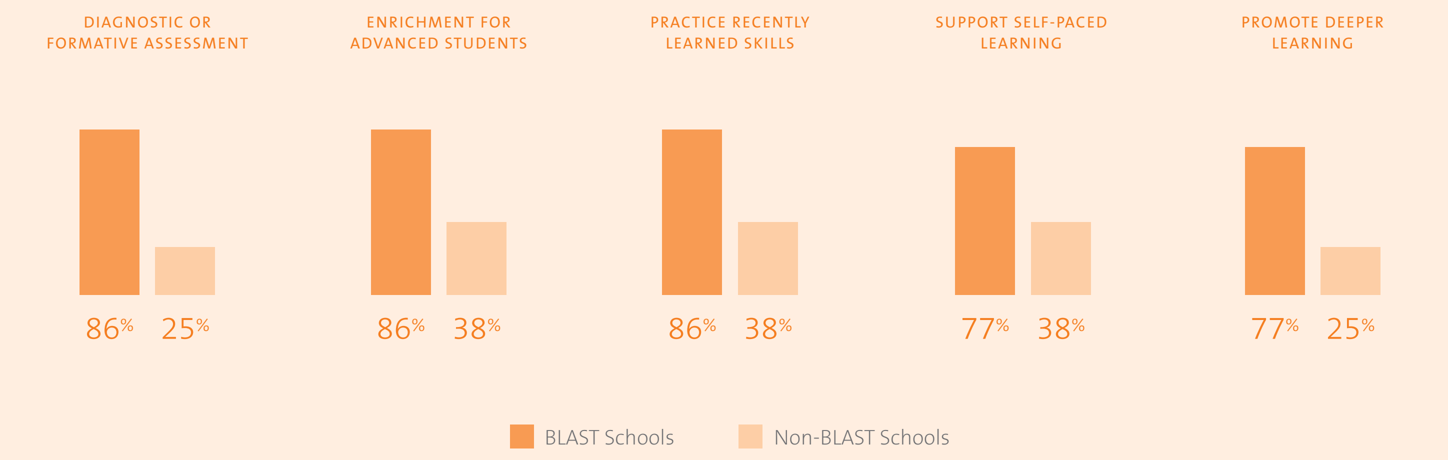
Organization of instruction. Based on teachers' responses to the study's survey, we found that ELA and math teachers in the BLAST schools spent more time providing opportunities for students to engage in small-group collaborative projects, small-group instruction, and, to a lesser extent, independent work compared to teachers in non-BLAST schools. The increased time teachers in the BLAST schools devoted to group work and independent practice was consistent with the activities' structure facilitated by the station rotation model within BLAST classrooms. However, in ELA classrooms, BLAST teachers reported that they provided fewer opportunities for students to engage in self-directed learning activities compared to non-BLAST teachers, and less time working with students one-on-one. This is in contrast to math instruction, where BLAST teachers reported that their students spent significantly more time in self-directed instructional activities than their peers in the non-BLAST schools.

The difference in the amount of self-directed learning opportunities reported for ELA and Math in BLAST schools is likely due to the significant differences in the amount of time students spent in online instruction in the two subject areas. Math students spent about seven times more time in online instruction than ELA students (approximately 60 hours on average in math across the entire school year compared to 8 hours for ELA).

Role of technology in instruction. Teachers implementing the BLAST model and those in the comparison group both reported using technology and web-based instruction to "personalize" the learning experience in similar and different ways. A majority of teachers (60% or more) in both the BLAST and non-BLAST schools reported that they used technology to support their instruction in the following ways: to meet the needs of different types of learners; to provide an additional way for students to access material; for remediation purposes; as a source of student learning data that can be used to inform instruction; and for test preparation. In addition, about half the teachers in both groups reported that they used technology-supported instruction to help introduce students to new concepts within a core lesson.

Although there were similarities in the roles that technology played in BLAST and non-BLAST schools, there were also some significant differences as well. BLAST teachers were much more likely than non-

ROLE OF TECHNOLOGY IN BLAST AND NON-BLAST SCHOOLS



BLAST teachers to report that they used technology for diagnostic or formative assessment (86% vs. 25%); for enrichment for advanced students (86% vs. 38%); to provide opportunities for students to practice recently learned skills (86% vs. 38%); to support self-paced learning (77% vs. 38%); and to promote deeper learning (77% vs. 25%).

It is also important to note that although non-BLAST teachers reported using technology to support their instruction, access to technology in non-BLAST schools and the frequency with which it was used in classrooms was limited compared to the BLAST schools. For example, only 20% of teachers in comparison schools reported that they had enough computers in their classroom for each student to have access to their own computer (compared to 91% of teachers in the BLAST schools). Also, more than half of the non-BLAST teachers indicated that students did not receive any online or computer-mediated instruction in ELA or math in a typical week.

INFRASTRUCTURE AND TECHNOLOGY

Unreliable connectivity and inadequate bandwidth were cited by teachers as significant technology-related factors impacting their ability to implement their blended learning models. A majority of BLAST teachers (95%) reported that either they or their students experienced technical problems during the year and that these problems were either moderate or significant barriers to effective use of the BLAST model. The most commonly cited technical issues included the reliability of their Internet connectivity and software, and insufficient bandwidth that impacted the use of some online programs.

Teachers and school administrators emphasized the importance of having on-site IT staff available to address technology issues when they arose. Both Burton and Simon had dedicated information technology (IT) support staff located on campus. Burton had on-site IT staff available throughout the school year, while Simon did not hire a dedicated IT support person until the end of the school year. Tennenbaum hired IT support by the hour on an as-needed basis. Teachers at Burton reported a limited number of disruptions to instruction due to network failures (no more than several days over the year) and attributed this to the responsiveness of its IT department. Several teachers at Simon reported that, prior to the hiring of their IT support technician, they regularly experienced problems with technology that impacted instruction. Many of the problems were relatively minor, such as computers being unable to play instructional videos because they did not have the latest version of Adobe Flash Player; thus, because IT support was not readily available, students would lose access to the online resource until the problem was fixed.

Michael & Susan Dell FOUNDATION

TRAINING AND SUPPORT

Teachers were not satisfied with the training they received to support their implementation of the BLAST model. Almost all of BLAST teachers (91%) participated in a training or orientation session directly related to their school's use of the BLAST or on the specific software programs supporting it. Of those who participated, 60% reported being dissatisfied with the quality of the training and more than 80% of teachers reported that insufficient training and lack of planning time had a "moderate" or "significant" impact on their ability to effectively implement the BLAST model. The primary reasons cited by those teachers who were dissatisfied with the training provided was that the training was too general and did not provide enough concrete examples they could use in their classrooms. The most beneficial training sessions reported by teachers were the ones led by fellow teachers in their own subject area, allowing these teachers to share ideas and strategies that were directly applicable to the content they were teaching. More than 90% of BLAST teachers reported spending some of their own time (at least an hour or more) getting acquainted with the online programs or planning for how to best integrate the BLAST model in their classrooms. Forty percent of teachers reported spending 20 hours or more. Teachers of science, social studies, and foreign language reported spending the most personal time preparing to implement the BLAST model.

SOFTWARE SELECTION, AVAILABILITY, AND DESIGN

Access to Compass Learning and Apex Learning courses varied by school. Tennenbaum students had earlier access to Compass Learning and Apex Learning courses than the other BLAST schools. According to Alliance leadership, Simon and Burton did not access Compass Learning until January, while students in Tennenbaum had access to these programs from the start of the school year. Several factors contributed to the delay of the roll out of Compass Learning in Simon and Burton, including a lack of funding at the start of the year to purchase a program subscription for the schools and the schools' decision to stagger the school-wide roll out of Compass Learning so teachers could first become familiar with the program and its features before committing to using it with their students. According to Alliance and school leadership, budgetary constraints also prevented the timely purchase of Apex Learning seat licenses in Burton and Simon, which limited students' access to Apex Learning courses in these schools during the school year (Burton students did not enroll in any Apex Learning courses during the 2011–12 school year). Also, since Tennenbaum had a dedicated learning lab for online coursework and Simon and Burton did not, Tennenbaum students had greater access to Apex Learning courses than students in the other schools. By the end of the school year, about 15% of all 9th and 10th graders at Tennenbaum enrolled in an Apex course in math (21 students of 152) and 11% in ELA. The rate of enrollments was similar in Simon (15% in math and 12% in ELA). Participation in Apex courses other than ELA and math courses was much more prevalent in Tennenbaum than Simon. More than 70% of Tennenbaum students enrolled in an Apex Learning course as an "elective" (e.g., AP Biology, Psychology, Spanish). In contrast only seven students in 9th and 10th grade in Simon enrolled in an elective class through Apex Learning.

Several challenges were reported by teachers that affected their ability to integrate online and teacher-led instruction. Two-thirds of teachers implementing the BLAST model reported that the lack of alignment between the online programs and the core curriculum and teacher-led instruction had "moderate" or "significant" impacts on their ability to effectively use blended learning with their students. Information collected during our interviews provided some additional details about this issue. For example, two ELA teachers spoke of the difficulty they had in identifying the relevant content in the some of the online programs to support their lessons. According to these teachers, the process was inefficient, time-consuming, and not always successful. In addition, a few BLAST history teachers using Compass Learning related that there was often a mismatch between the time allocated for the independent workstation—40 minutes—and students' ability to successfully complete a "lesson" assigned by the teacher within Compass Learning in that time. As a result, teachers did not always use the Compass Learning program when they covered a particular topic and instead identified other online activities that students could complete in the allocated time.

Software quality varied across subject areas. BLAST teachers and school administrators tended to be more satisfied with the content coverage and quality of software programs developed for math than they were for other subjects, including ELA, science, social studies, and foreign language. Alliance leadership also expressed concerns about the ability of available online programs for ELA to adequately support writing instruction. In particular, the leadership was concerned that assignments that require students to generate short written responses were insufficient to prepare students for college-level writing.

Availability of online programs for the sciences, social sciences, and humanities was limited. One of the more common challenges reported by teachers in the social sciences related to the availability of software. Specifically, many foreign language, science, and social studies teachers reported that they did not have access to an online instructional program in their content area, and instead had to compile various digital resources from the Web to implement the online and computer-mediated components of the BLAST model. This issue was partially resolved when teachers at Simon and Burton received access to Compass Learning.

USE OF DATA TO INFORM INSTRUCTION

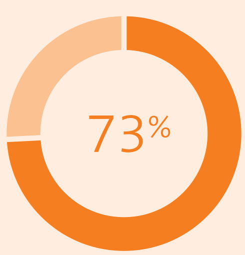
A majority of teachers utilized the reports of student progress and performance captured by the online programs. Almost three-quarters of the teachers implementing the BLAST model (73%) reported that they reviewed student performance data provided by the online programs, and they did so regularly. All but one teacher reported that they reviewed student data at least once a week and almost a third of teachers reviewed student reports on a daily or near daily basis. Among teachers who never reviewed student data (about 1 in 4 teachers), the most cited reasons for not reviewing data included the data not being informative or relevant and that they relied more on information outside of the online programs for student evaluations.

Teachers in subject areas other than math and ELA also reported that they didn't have access to online programs that provided reports on student progress within the programs.

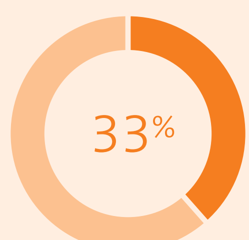
Seventy percent of teachers who reviewed the student reports from online programs reported they found the student performance data somewhat (41%) or very useful (29%) for informing their instruction. Math teachers' reports of the usefulness of the student performance data captured by the software program was more favorable than other subject-area teachers, particularly when compared to those teaching science, social studies, or a foreign language. In general, teachers implementing the BLAST model reported that reviewing student performance data from the online programs had allowed them to monitor and diagnose student understandings of key concepts at the individual and group level; provide students with feedback on their performance; and modify their plans for future lessons and instructional activities.

The use of multiple online programs posed some unique challenges to the utilization of data from the online programs. The BLAST schools adopted multiple online programs from different vendors that used different metrics and formats for displaying and outputting student data. As a result one major challenge to the effective use

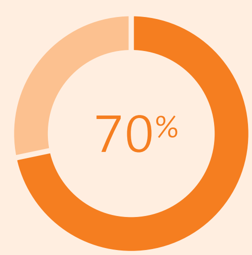
USE OF DATA TO INFORM INSTRUCTION



NUMBER OF TEACHERS WHO REVIEWED STUDENT PERFORMANCE DATA PROVIDED BY THE ONLINE PROGRAMS



NUMBER OF TEACHERS WHO REVIEWED STUDENT REPORTS ON A DAILY OR NEAR DAILY BASIS



NUMBER OF TEACHERS WHO REVIEWED STUDENT REPORTS FOUND THE DATA SOMEWHAT (41%) OR VERY USEFUL (29%)

of student performance data provided by the online programs to inform instruction was the difficulty administrators and teachers had in tracking and interpreting student progress across multiple and incompatible program reports. The process was extremely burdensome and inefficient. To resolve this issue, Alliance contracted with Education Elements during the 2011–12 school year to develop a data dashboard that would integrate the student data from the multiple online systems with schools' student data information system. A working dashboard was not yet available to administrators and teachers as of the research team's spring visit.

SATISFACTION WITH THE BLAST MODEL AND PERCEIVED IMPACTS ON STUDENTS

A majority of teachers were not satisfied with the implementation of the BLAST model in the school year 2011–12. Just over half of teachers (55%) surveyed reported that they would not recommend the use of the BLAST model to other teachers, and 60% of teachers disagreed with the statement that student learning had improved as a result of the BLAST model. Several different factors may be contributing to teacher dissatisfaction. For example, some BLAST teachers expressed challenges with classroom management in implementing the station rotation model with high school students. This issue was particularly salient in Tennenbaum, which had 48 students per classrooms. Other teachers, particularly those in subject areas other than math, also indicated that the three-station model (direct instruction, independent, and collaboration) was less conducive to the content they were teaching. Several teachers expressed frustration with the fact that they could no longer use lesson plans they had found to be successful in the past because the lessons were not compatible with a station rotation model.

Teachers' perceptions of the impacts of the BLAST model on student outcomes were mixed. Whereas two-thirds of BLAST teachers (68%) reported on the survey that they agreed with the statement that the BLAST model helped students take ownership over their learning, less than half of the teachers agreed that the model helped improve student learning (41%) and that it met the learning needs of their students (41%). ELA teachers were the least positive in their responses regarding the BLAST model improving students' learning and understanding. Teachers also reported that the BLAST model might not be benefiting all students equally. While a majority of teachers implementing the BLAST model reported that the model was "very effective" for students whose academic work was ahead of most students their age (59%), and at least "somewhat" or "very effective" for students working at grade-level (91%), a majority of teachers reported that the model was "not at all effective" for students whose academic work was behind most students their age (59%), or for special education students (59%).

Students' computer literacy skills were an important factor in whether students could fully leverage the beneFits of blended learning. Teachers and school administrators acknowledged that the BLAST model placed great demands on students' ability to self-direct their learning and use technology as a tool for learning. They related that since many students in the BLAST schools lacked basic computer literacy and skills, this hindered students' ability to take full advantage of the computer-mediated instruction. More than 60% of teachers reported the students' lack of computer skills had a "moderate" or "significant" impact on their ability to effectively implement the BLAST model. Several teachers across schools and subject areas reported that in the future they would spend more time at the beginning of the school year reviewing basic computer, navigation, and software troubleshooting skills with students before implementing the station rotation model in their classrooms.

Access to nonacademic Internet content impacted students' engagement in online instructional environments. Teachers reported, through the survey and interviews, that a lack of student engagement within the online instructional programs may have negatively impacted the effectiveness of the BLAST model in some classrooms. Only half the teachers surveyed agreed with the statement that students are "highly engaged" while using the online programs (50%) and almost two-thirds of teachers reported (64%) that students' resistance to using some of the software programs had a "moderate" or "significant" impact on their ability to effectively implement the BLAST model. Off-task behavior, facilitated by students' access to the Internet, was a concern for most teachers and was observed during the site visits by researchers. Having access within the classroom to social media websites such as Facebook and GChat and other popular websites such as YouTube and iTunes clearly distracted some students from productive work in the online programs. One ELA teacher interviewed described this issue as a "constant battle between [teachers], Facebook, YouTube, and iTunes."

To address issues of student engagement and off-task behaviors, many teachers stressed the importance of effective classroom management strategies and establishment of a productive classroom culture. The

BLAST schools devised a number of different incentive and motivational strategies to keep students engaged and on-task while working on the online programs within the independent and collaboration workstations. Specific strategies used by teachers included making progress and performance in the online instructional programs a portion of the class grade; the use of fellow students as "station leaders" to promote peer support and encouragement; and, in some classrooms, allowing students to listen to music while working on the online programs. Support for allowing students to listen to music while completing their online work varied across schools and classrooms. While some teachers used music as a reward for students who were making adequate progress, others reported they stopped the practice after determining it was more of a distraction than an incentive.

Teachers also debated the need for schools to adopt more effective Internet filters to limit off-task Internet activities. In general, teachers across the BLAST schools tended to be dissatisfied with the existing filters used by the schools because they blocked not only student access to inappropriate Internet content, but also teachers' access to online instructional materials. Teachers expressed a desire for more sophisticated filters that would allow them access to materials while still blocking students from unauthorized sites such as Facebook or Gchat. However, there were some differences of opinion among teachers regarding the need for filters. While some teachers advocated for improved filters, others believed that, instead, more time should be spent working with students on how to use the Internet responsibly during class time and as a tool for learning.

FINDINGS FROM THE INSTRUCTIONAL SYSTEM LOG DATA

To capture the intensity with which students used the different core online learning programs, we computed the number of minutes that a student logged over the course of the school year for each program. Data was analyzed from Revolution K12 and Compass Learning (ELA and math instruction) as well as Achieve3000 (ELA only). Burton and Simon did not start using Compass Learning until January 2012. Tables 6 and 7 provide descriptive statistics for the number of minutes

TABLE 6: DESCRIPTIVE STATISTICS FOR INTENSITY OF USE OF ONLINE PROGRAMS, ELA (MINUTES)

	REVOLUTION K12		ACHIEVE3000		COMPASS LEARNING		TOTAL
	N (OVERALL)	MEDIAN (MIN/MAX)	N (OVERALL)	MEDIAN (MIN/MAX)	N (OVERALL)	MEDIAN (MIN/MAX)	MEDIAN (MIN/MAX)
9TH GRADE							
TFTA[A]	-	-	76 (76)	63 (5/146)	69 (76)	54 (2/259)	111 (9/352)
Burton	141 (154)	108 (61/349)	152 (154)	55 (1/114)	49 (154)	84 (15/235)	176 (32/588)
Simon	-	-	168 (173)	57 (4/185)	116 (173)	31 (5/64)	81 (4/200)
10TH GRADE							
TFTA	65 (76)	228 (71/1,289)	76 (76)	74 (14/214)	62 (76)	56 (9/366)	366 (34/1,564)
Burton	134 (149)	510 (75/845)	127 (149)	76 (30/176)	-	-	543 (30/906)
Simon	16 (123)	109 (73/194)	120 (123)	42 (1/165)	5 (123)	23 (5/131)	45 (1/235)

TABLE 7. DESCRIPTIVE STATISTICS FOR INTENSITY OF USE OF ONLINE PROGRAMS, MATH (MINUTES)

	REVOLUTION K12		ACHIEVE3000		TOTAL
	N (OVERALL)	MEDIAN (MIN/MAX)	N (OVERALL)	MEDIAN (MIN/MAX)	MEDIAN (MIN/MAX)
9TH GRADE					
TFTA[A]	71 (76)	1,171 (152/5,029)	70 (76)	681 (102/3,020)	1,820 (152/7,335)
Burton	147 (154)	2,506 (109/5,476)	97 (154)	41 (4/145)	2,489 (58/548)
Simon	169 (173)	1,287 (248/3,294)	164 (173)	655 (23/1,915)	1,984 (336/452)
10TH GRADE					
TFTA	73 (76)	2,386 (545/4,827)	73 (76)	1,043 (31/5,171)	3,596 (159/7,460)
Burton	131 (149)	1,863 396/3,977	136 (149)	97 (14/318)	1912 (14/422)
Simon	99 (123)	648 (72/1,609)	113 (123)	742 (30/1,950)	1,343 (72/2,686)

that a student spent on each of the core online programs broken out by school, grade level, and subject area. The findings from the analysis of online program data are summarized next.

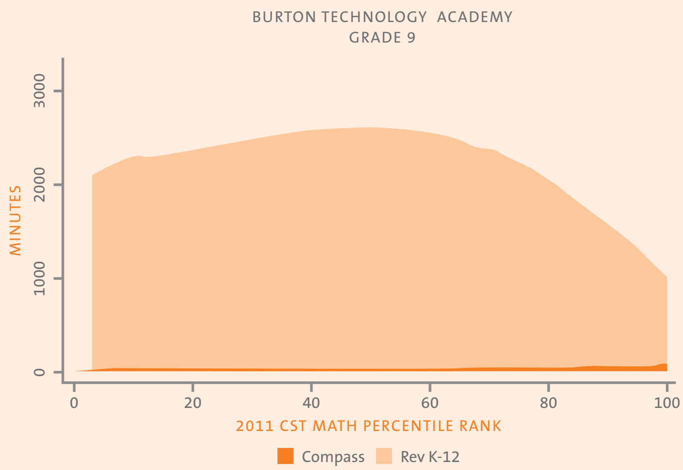
Online instruction was more frequently used to support math instruction than ELA instruction. Students logged nearly seven times more minutes on the core online math programs compared to the core programs used to support ELA instruction (see Tables 6 and 7). Median student use in total minutes of online math instruction logged across both 9th and 10th grade ranged from 1,300 minutes (Burton, 10th grade) to more than 3,600 (Tennenbaum, 10th grade) over the course of the school year across Compass Learning and Revolution K12 (representing about 34 to 95 minutes per week in a 38-week school year). In comparison, for online ELA instruction, median student use in 10th grade ranged from 45 minutes (Simon, 10th grade) to almost 400 minutes (Tennenbaum, 10th grade) during the school year (a few minutes to 14 minutes per week) and 80 to 175 minutes in 9th grade (a few minutes per week).

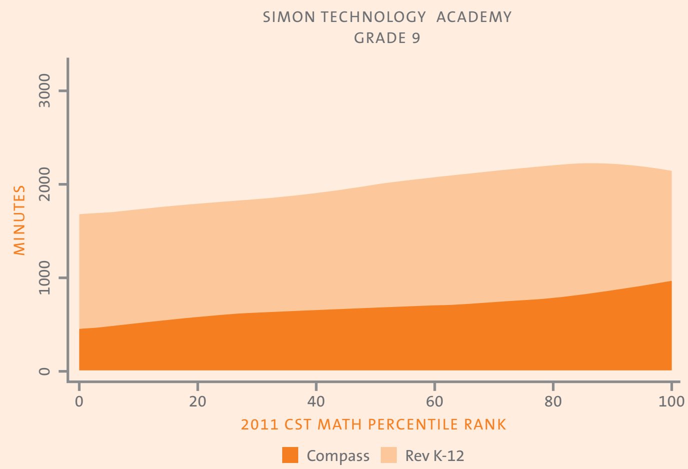
Use of online programs for math was differentiated by students' prior achievement levels. Figure 1 shows how time spent by students in 9th grade on the different core online instructional programs in math varied by students' prior academic performance. (A similar analysis was not possible for 10th grade due to the lack of availability of prior CST scores for Tennenbaum.) In general, there was evidence that the overall time spent in online math instruction using the core programs differed for 9th-grade students with different academic needs in two of the three schools (Tennenbaum and Burton), as students in these schools with the greatest needs spent more time using the online programs.

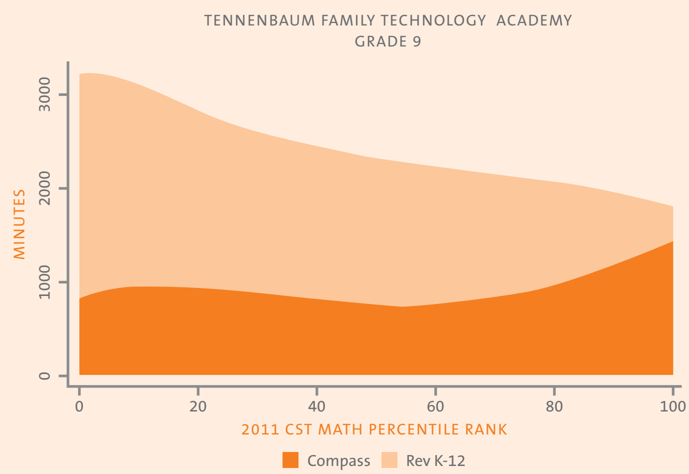
To illustrate these trends, Figure 1 illustrates how the number of minutes that students spent on each core online math program varied by students' prior achievement scores and by school (spring 2011 California Standards Test score). For example, in Burton there was a significant drop-off in the use of Revolution K12 and overall time in online instruction for the core online programs with increasing prior achievement (about a 50% drop when comparing students scoring

[A] *TFTA = Tennenbaum Family Technology Academy*

FIGURE 1. DISTRIBUTION OF TIME OF USE BY PRIOR ACHIEVEMENT BY SCHOOL, MATH. (NOTE: PERCENTILES WERE GENERATED SEPARATELY FOR EACH SCHOOL. A STUDENT AT THE 50TH PERCENTILE WITHIN BURTON MAY NOT BE AT THE 50TH PERCENTILE IN TENNENBAUM. THE PERCENTILES ARE NOT EQUIVALENT TO STATE NORMS.)







between the 60th and 100th percentile on the spring 2011 CST). In Tennenbaum, we also saw significant decrease in Revolution K12 and overall time in online instruction with increasing prior academic performance (about a 30% drop between students in the lowest and top quartile of prior academic performance). In contrast, we found a slight increase in overall use (about 20%) by students with higher prior achievement in Simon.

The differences in time logged on the core online systems for math between students of different abilities in Tennenbaum and Burton may have been a function of the extra opportunities that were available to lower-achieving students to access online instruction to support their needs. Students with lower prior achievement may have been given additional opportunities to receive online instruction, particularly through the use of Revolution K12, during after-school support sessions and during the Saturday Academy to help fill in gaps in foundational Algebra I skills.

Impact Analyses

IMPACT DESIGN OVERVIEW

A quasi-experimental design was implemented to explore the relationship between adoption of the BLAST model and gains in student learning compared to similar schools not using the BLAST model. The intent of the design was to understand what impact the BLAST model may have had on students' academic achievement relative to instructional environments that the students may have been exposed to if their schools had not adopted the BLAST model. The design compared the learning outcomes for students enrolled in the three Alliance-BLAST schools (Burton, Simon, and Tennenbaum) in the school year 2011–12 to comparison students attending one of three other high schools in the Alliance charter school network (Ouchi, College Ready Academy #5, Health Services Academy) that had not yet adopted the BLAST model. Scores on the spring 2012 achievement assessments were compared between the three BLAST schools and three Alliance comparison schools, controlling for prior student achievement scores (additional details on the analytical models used are included in the report's Technical Appendix).

The comparison schools were selected by Alliance leadership for their similarities to the three BLAST schools, including the schools' student composition, leadership, school culture, and prior academic performance. As described above, the comparison schools were considered high-performing schools based on a comparison of state achievement scores to other schools in the state serving similar student populations (the comparison schools ranked in the top 20% of similar schools). In contrast, one of the BLAST schools, Simon, ranked below the average high school in the state serving similar student populations and the other school with similar school ranking, Burton, scored in the top 10% (Tennenbaum opened in 2011 and therefore did not have a similar school ranking for the 2010–11 school year).

To be eligible to be included in the impact analyses for a particular outcome measure, 75% or more of the eligible students in a school and grade level had to have "complete" data—scores on both the outcome and prior achievement measure. As a result, Tennenbaum was excluded from two of the analyses—an analysis of scores on the 9th-grade math CST and 10th-grade Alliance ELA benchmark assessment—due to a high level of missing data for prior achievement scores. In addition, several other outcome measures could not be included in the analyses because of the high level of noncomplete cases across both BLAST and non-BLAST schools including an internal math benchmark (9th and 10th grade), Math California High School Exit Exam (10th), and 10th-grade math CST.[12]

Descriptive statistics for the BLAST and non-BLAST student samples included in the impact analyses appear in Table A1 in the Appendix for this profile. The most notable differences in student demographics between the BLAST and non-BLAST schools were for eligibility for the federal subsidized lunch program in 9th grade (a measure of family poverty) and in 10th grade the percentage of students identified as Black or Hispanic. Specifically, for the 9th-grade analytic sample, on average BLAST schools had fewer students who were eligible for the federal lunch program compared to non-BLAST schools (82% compared to 92%) although the numbers were quite high relative to national averages. In the 10th-grade analytic sample, BLAST schools on average included fewer Black students compared to non-BLAST schools (9% compared to 22%) but more Hispanic students (89% compared to 78%).

Table A2 in the Appendix for this profile shows the equivalence between the BLAST and non-BLAST schools on prior achievement scores for the student samples included in the impact analyses. There were no statistically significant differences in prior achievement between BLAST and non-BLAST students in the analysis sample.

Limitations of design. Although the comparison schools may be similar in many ways to the schools implementing the BLAST model during the 2011–12 school year, there may have been important differences between the BLAST and the comparison schools (e.g., differences in curriculum, academic culture, school leadership, and teaching staff) that may explain differences between the schools on spring test scores that are totally unrelated to the use the adoption of the BLAST model. Thus, using this design we could not isolate the effect of the introduction of BLAST instruction from other key differences between schools that are likely to influence student achievement.

[12] *Because the prior achievement scores for the analysis of outcomes for 10th-grade students had to come from the 2009–10 school year (the year prior to the year Simon and Burton first piloted the BLAST model), the level of missing prior achievement data was greater for 10th grade than it was for 9th grade.*

Table 8 shows the results of the impact analyses comparing students enrolled in BLAST schools to students in the comparison schools controlling for differences in prior achievement between students in the two groups of schools. Results are shown for student performance on Alliance's benchmark assessments in ELA and math, and 9th grade state achievement scores (CST). Tenth-grade CST were not analyzed due to high levels of missing prior achievement scores for this group of students.[13] High levels of missing data was also the reason the inclusion of Tennenbaum students in the impact analyses was limited to the analysis of 9th-grade ELA CST scores and scores on Alliance's benchmark assessments.

IMPACT FINDINGS

We found no evidence of greater student learning gains in math or ELA associated with attending Alliance-BLAST schools relative to three comparison schools in the Alliance network. The analyses revealed statistically significant differences in spring 2012 test scores (p-value < 0.05) between students in the BLAST and comparison schools for several different outcome measures. In each case, after controlling for prior achievement in the models, the differences favored students in the comparison schools. As shown in Table 8, students attending Alliance-BLAST schools scored lower compared to students in the comparison

TABLE 8: OVERALL EFFECTS ON TEST SCORES (BLAST VS. COMPARISON SCHOOLS, SY 2011–12)

OUTCOMES	TREATMENT			COMPARISON					
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	MODEL-ADJUSTED MEAN	P
GRADE 9									
2012 ELA Benchmark	313	3.1	0.8	405	3.2	0.6	-0.1	-0.17	0.037
2012 CST ELA	316	325.0	51.2	405	339.1	47.0	-14.1	-0.29	< 0.001
2012 CST Algebra[A]	212	310.0	61.7	313	339.4	63.7	-29.4	-0.47	0.026
GRADE 10									
2012 ELA Benchmark[A]	171	3.1	0.7	264	3.2	0.6	-0.1	-0.21	0.286

Note: Model adjusted means are expected values after controlling for differences in prior achievement; estimated impact is the difference in model adjusted means (BLAST–Comparison); effect size = estimated impact from model/pooled standard deviation (SD) between BLAST and comparison schools.

[A] *Sample does not include students in the Tennenbaum Family Academy.*

TABLE 9: EFFECT SIZES AND CORRESPONDING MEAN 9TH-GRADE SPRING CST TEST SCORES AND CST PERFORMANCE LEVELS BY EXPERIMENTAL GROUP*

OUTCOMES	ADJUSTED MEAN CST SCORE				BLAST (% AT PERFORMANCE LEVELS)[B]					NON-BLAST (% AT PERFORMANCE LEVELS)				
	EFFECT SIZE	BLAST	NON-BLAST	DIFFERENT CST SCALE POINTS	FBB	BB	B	P	A	FBB	BB	B	P	A
SUBJECT														
Algebra 1	-.47*	310.0	339.4	29.4	19	34	22	21	4	7	24	26	33	10
English/Language Arts	-.29*	325.1	339.2	14.1	12	20	32	27	9	6	15	40	28	10

[A] *Note: The mean scores and differences are "adjusted" model-based estimates. However, for the descriptive percentages for performance levels there are no statistical controls for prior achievement.*

[B] *Performance levels: FBB = Far Below Basic, BB = Below Basic, B = Basic, P = Proficient, A = Advanced*

Contrast between Treatment and Comparison statistically significant, p < .05

[13] *This was due to the use of spring 2010 scores for the prior achievement measure, since Simon and Burton first adopted the BLAST in fall 2010 when the current 10th-grade cohort was in 9th grade.*

TABLE 10: EFFECTS ON TEST SCORES BY SCHOOLS (BLAST VS. COMPARISON SCHOOLS, SY 2011–12)

OUTCOMES	TREATMENT			COMPARISON					
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
GRADE 9									
2012 ELA Benchmark									
Tennenbaum	56	3.0	0.7	405	3.2	0.6	-.02	-0.29	0.357
Simon	125	3.2	0.9	405	3.2	0.6	0.0	-0.06	0.773
Burton	132	3.0	0.7	405	3.2	0.6	-0.2	-0.34	0.214
2012 CST ELA									
Tennenbaum	56	330.5	49.7	405	338.6	47.0	-8.1	-0.17	0.285
Simon	128	322.5	50.2	405	338.6	47.0	-16.1	-0.34	0.001
Burton	132	328.5	51.3	405	338.6	47.0	-10.1	-0.21	0.098
2012 CST Algebra I[A]									
Simon	108	330.5	49.7	313	328.5	51.3	-55.8	-0.94	0.000
Burton	104	322.5	50.2	313	328.5	51.3	2.8	0.04	0.693
GRADE 10									
2012 ELA Benchmark									
Simon	63	3.2	0.7	264	3.2	0.6	0.0	0.02	0.967
Burton	108	3.0	0.7	264	3.2	0.6	-0.3	-0.40	0.356

Note: Model adjusted means are expected values after controlling for differences in prior achievement; estimated impact is the difference in model adjusted means (BLASTComparison); effect size = estimated impact from model/pooled standard deviation (SD) between BLAST and comparison schools.

[A] *Sample does not include students in the Tennenbaum Family Academy.*

schools on the spring 9th-grade state achievement test in Algebra I (Simon and Burton only) and ELA; Alliance's own 9th-grade ELA internal benchmark assessment; and the 10th-grade CAHSEE for ELA.

9th-grade Algebra I CST (Table 8 and Table 9). As shown in Table 8, the average 9th-grade student in two of the three BLAST schools with sufficient data—Simon and Burton—underperformed students in the comparison schools on both the ELA and math CST (effect size = -0.47). The difference represents students in the BLAST schools scoring lower by 29 scale points or the median student in the comparison schools dropping 18 percentiles if they had attended one of the BLAST schools (moving from the 50th to the 32nd percentile). Forty-seven percent of students in Simon and Burton scored at or above basic proficiency on the CST compared to 69% of students in the comparison schools (see Table 9).

Variation in effects on Algebra I scores across schools (Table 10). Separate models were run comparing spring 2012 Algebra I scores for students in Simon and Burton to scores for students in the non-BLAST schools controlling for differences in student prior achievement. We found no difference in test scores between spring test scores for BLAST and comparison students for Burton (effect size = +0.04; a difference of 3 scale points), while a large negative and statistically significant effect was found for Simon (effect size = -0.93; -56 scale points). Clearly, the results for Simon drove the negative difference in spring Algebra I CST scores between BLAST and non-BLAST schools as reported above.

9th- and 10th-grade ELA (Table 8 and Table 9). Ninth-grade students in BLAST schools on average also scored lower on the study's ELA outcome measures compared to 9th graders in the comparison schools (see Table 8). The differences were statistically significant (p-value < 0.05). All three BLAST schools were included in the analysis of spring ELA CST and the 9th-grade internal benchmark assessment, while the analysis of the 10th-grade internal benchmark assessment was restricted to Simon and Burton. Ninth-grade students in BLAST schools on average scored lower on both an internal benchmark assessment (effect size = -0.17 or a tenth of a point difference on a 4 point-scale) and the ELA CST (effect size = -0.29 or

Michael & Susan Dell FOUNDATION

14 scale points lower). We found no statistically significant differences between students in BLAST schools and their comparisons for the 10th-grade ELA internal benchmark. The size of the difference in CST scores (effect size = -0.29) represents the median student in the comparison schools dropping 11 percentiles if that student had attended a BLAST school (moving from the 50th to the 39th percentile). Sixty-eight percent of students in Alliance's BLAST schools scored at or above basic proficiency on the CST compared to 78% of students in the comparison schools (Table 9).

Variation in effects on ELA scores across schools (Table 10). We also examined the extent to which the effects of the BLAST model on ELA test scores reported in Table 8 varied by BLAST school. Similar to our approach for the analysis of individual school effects for math, separate models were run comparing spring scores for students in each of the BLAST schools to scores for students in the non-BLAST schools, controlling for differences in student prior achievement. For the analysis of 9th-grade ELA scores, we found no evidence that an individual school was disproportionately contributing to the negative effect reported above. Effect sizes associated with the individual schools ranged from -0.17 (a difference of 8 scale points) in Tennenbaum to -0.21 and -0.33 in Burton and Simon, respectively (or a difference of 10 to 16 scale points). The effect was statistically significant (p-value < .05) for Simon only. When the outcome was the 9th-grade ELA internal benchmark, we found that the negative effects of BLAST reported above across all BLAST schools tended to be greater for Tennenbaum and Burton (effect size = -0.29 and -0.33, respectively) than for Simon (effect size = -0.06). None of these individual effects were statistically significant on their own. Finally, for scores on the 10th-grade ELA internal benchmark, an outcome measure for which we found no statistically significant effect when comparing all BLAST and non-BLAST students combined across schools, we found evidence suggesting that the effects varied by school with essentially no difference estimated between Simon and the non-BLAST schools (+0.02) and a negative effect estimated for Burton (-0.40). None of these individual effects were statistically significant.

Subgroup analyses. We also analyzed the extent to which the differences in test scores between BLAST and non-BLAST students differed by gender and student prior achievement (see Table 11 and Table 12). We found no

TABLE 11. EFFECTS ON TEST SCORES, BY GENDER (BLAST VS. COMPARISON SCHOOLS, SY 2011–12)

OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
GRADE 9									
2012 ELA Benchmark									
Females	160	3.0	0.7	213	3.2	0.6	-0.2	-0.26	0.007
Males	153	3.0	0.8	192	3.2	0.6	0.0	-0.07	0.471
2012 CST ELA									
Females	162	325.8	49.3	213	339.4	47.8	-13.7	-0.28	<0.001
Males	154	324.5	53.1	192	338.9	46.2	-14.5	-0.29	<0.001
2012 CST Algebra I[A]									
Females	110	314.7	66.5	159	341.3	60.5	-26.6	-0.42	0.052
Males	102	304.7	53.6	154	337.5	67.1	-32.7	-0.52	0.017
GRADE 10									
2012 ELA Benchmark[A]									
Females	84	3.1	0.6	137	3.2	0.6	-.02	-0.27	0.179
Males	87	3.1	0.7	127	3.2	0.6	-.01	-0.16	0.417

Note: Model adjusted means are expected values after controlling for differences in prior achievement; estimated impact is the difference in model adjusted means (BLAST–Comparison); effect size = estimated impact from model/pooled standard deviation (SD) between BLAST and comparison schools.

[A] *Sample does not include students in the Tennenbaum Family Academy.*

TABLE 12: EFFECTS ON TEST SCORE BY STUDENT PRIOR ACHIEVEMENT (BLAST VS. COMPARISON SCHOOLS, SY 2011–12)									
OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
GRADE 9									
2012 ELA Benchmark									
High Prior	156	3.2	0.7	195	3.3	0.5	-0.1	-0.11	0.264
Low Prior	157	3.0	0.6	210	3.1	0.5	-0.1	-0.21	0.027
2012 CST ELA									
High Prior	157	335.7	35.3	195	345.1	38.0	-9.4	-0.19	0.005
Low Prior	159	315.4	33.8	210	333.2	35.2	-17.8	-0.36	<0.001
2012 CST Algebra I[A]									
High Prior	103	316.0	62.5	153	346.2	63.6	-30.3	-0.48	0.032
Low Prior	109	304.9	37.5	160	332.4	51.1	-27.5	-0.44	0.051
GRADE 10									
2012 ELA Benchmark[A]									
High Prior	81	3.2	0.5	137	3.3	0.5	-.0.1	-0.11	0.600
Low Prior	90	3.0	0.5	127	3.2	0.5	-0.2	-0.31	0.123

Note: Model adjusted means are expected values after controlling for differences in prior achievement; estimated impact is the difference in model adjusted means (BLAST–Comparison); effect size = estimated impact from model/pooled standard deviation (SD) between BLAST and comparison schools.

[A] Sample does not include students in the Tennenbaum Family Academy.

statistically significant differences in the effects of attending BLAST schools between males and females or students with different levels of prior achievement.

ACHIEVEMENT GAP ANALYSIS

In addition to estimating the impact of the BLAST model by comparing students' test scores between schools implementing BLAST and a set of comparison schools that were not, we also examined **the extent to which BLAST schools narrowed the achievement gap between low-income and high-income schools on state test scores.** The results of the analyses of the achievement gap for BLAST schools are presented in Tables 14 and Figures 1 and 2 for spring 2012 9th-grade ELA test scores only. The results for 10th grade are similar. We did not conduct an achievement gap analysis for CST math scores because of the nature of math course-taking in high school that makes comparisons between high-income and low-income high schools problematic.[14]

Table 13 shows the sample size for the different groups of schools included in the achievement gap analysis for Alliance. We define high-income schools as other public high schools in California with 10% or fewer of its students qualifying for the federal free or reduced-price lunch program. Low-income schools are those schools in the state whose proportion of students who qualify for free or reduced-price lunch is within plus or minus 5% of the Alliance's BLAST and non-BLAST schools. We examined the high-income/low-income achievement gap for both the BLAST schools as well as the non-BLAST comparisons to determine the extent to which the BLAST schools may be closing the gap of and beyond other schools in the Alliance network.

We used two different approaches to estimating the reduction in the achievement gap in ELA test scores between high-income and BLAST schools in state achievement scores: a comparison of performance levels and scale point scores (for details on the achievement gap analyses,

[14] *In California high schools, the state math assessment taken by a student depends on which math course a student is enrolled in during the school year. If students in high-income schools are more likely to take more advanced math courses such as Algebra II or Geometry in 9th and 10th grade compared to students in low-income schools, the majority of which may be enrolled in Algebra I, then any comparison of Algebra I scores between high- and low-income schools may represent a comparison between the lowest performing students in the high-income schools to some of the better performing students in the low-income schools. As a result any differences found between the groups may be biased favoring the low-income schools and thus difficult to interpret.*

see description above in the Analytical Models section on page 17). As a result of applying these approaches we estimated that the BLAST schools reduced the income achievement gap in performance on the state 9th-grade ELA assessment by about 3% (The results are similar for 10th grade.) However, there is little evidence that Alliance-BLAST schools closed the achievement gap relative to the three Alliance comparison schools. The results are summarized in Table 14 and in Figures 2 and 3.

Comparison of performance levels. As shown in Table 14 and Figure 2, the current difference, or gap, between high-income and low-income schools in California in the percentage of students scoring at or above proficient on the ELA California Standards Test is 33.5%. When we examine the achievement gap between Alliance-BLAST schools and high-income high schools in California, the gap is reduced to 30.5% for ELA or a 3% reduction in the overall gap. However, when we compare this reduction to the reduction of the gap for the Alliance comparison schools (from 33.5% to 31.5% or a reduction of 2%), we see that the reduction of the gap for the BLAST schools is only 1% more than reduction for the Alliance comparison schools.

Comparison of scale points. When the unit of comparison is differences in the average scale point scores (see Table 14 and Figure 3), we found that the overall reduction in the income-based achievement gap for students in the BLAST schools is less than the reduction in the gap for students in the Alliance comparison schools. The estimated reduction in the income-based achievement gap associated with enrollment in Alliance-BLAST schools is 1.6 scale points in ELA (a reduction from 53.74 points to 52.12 points or 3%). In contrast the estimated reduction in the income-based achievement gap associated with enrollment in Alliance comparison schools is 8.6 scale points in ELA (a reduction from 53.74 points to 45.25 points) or a reduction of 16% of the overall income-based gap.

TABLE 13: SAMPLE SIZE FOR THE ACHIEVEMENT GAP ANALYSIS

	ELA
Alliance-BLAST Schools	3
Alliance non-BLAST comparison school	3
Low-income comparison schools	211
High-income comparison schools	146

TABLE 14. SUMMARY OF ACHIEVEMENT GAP RESULTS BY COMPARATIVE MEASURES (CST ELA)

COMPARISONS	ELA
% OF STUDENTS SCORING AT OR ABOVE PROFICIENT	
High- vs. Low-income	33.5%
High-income vs. BLAST schools	30.5%
[High-income vs. non-BLAST schools]	[31.5%]
% Reduction in Achievement Gap	**3%**
[% Reduction in non-BLAST schools]	[2%]
SCALE POINTS	
High- vs. Low-income	54 pts
High-income vs. BLAST schools	52 pts
[High-income vs. non-BLAST schools]	[45 pts]
% Reduction in Achievement Gap	**3%**
[% Reduction in non-BLAST schools]	16%

FIGURE 2: DIFFERENCE IN PERCENTAGE OF STUDENTS AT OR ABOVE PROFICIENT BETWEEN ALLIANCE-BLAST SCHOOLS, ALLIANCE COMPARISON SCHOOLS, AND HIGH- AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA.

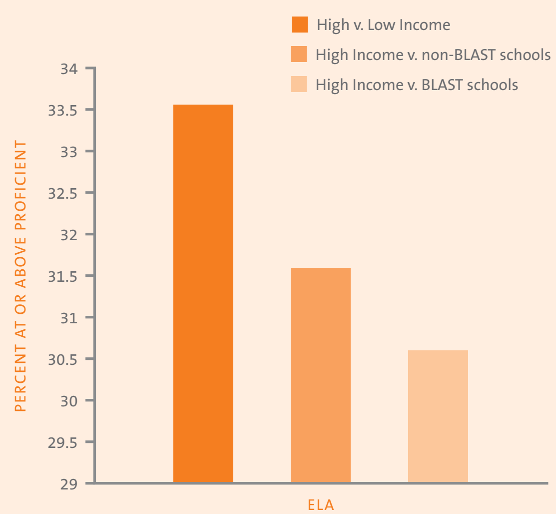
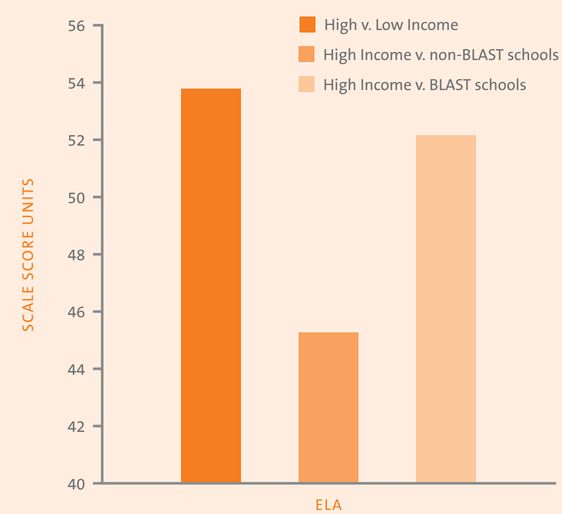


FIGURE 3. DIFFERENCE IN ELA AND MATH SCALE SCORE POINTS BETWEEN ALLIANCE-BLAST SCHOOLS, ALLIANCE COMPARISON SCHOOLS, AND HIGH- AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA.



Michael & Susan Dell FOUNDATION

Appendix: Alliance College-Ready Public Schools

TABLE A1: BASELINE DESCRIPTIVE STATISTICS: STUDENTS IN THE ANALYTIC SAMPLE, BY PERCENT

	(BLAST) ALLIANCE TENNENBAUM FAMILY TECHNOLOGY HIGH SCHOOL (%)	(BLAST) SIMON TECHNOLOGY ACADEMY HIGH SCHOOL (%)	(BLAST) BURTON TECHNOLOGY ACADEMY HIGH SCHOOL (%)	COLLEGE-READY ACADEMY HIGH SCHOOL #5 (%)	HEALTH SERVICES ACADEMY HIGH SCHOOL (%)	WILLIAM AND CAROL OUCHI HIGH SCHOOL (%)
9TH GRADE						
English Language Status						
English Language Learner (ELL)	31	26	20	25	26	29
Not ELL	69	74	80	75	74	71
Gender						
Female	41	51	55	44	57	57
Male	59	49	45	56	43	43
Free/Reduced-Price Lunch (FRPL)						
FRPL Eligible	69	77	99	94	87	96
Not FRPL Eligible	31	23	1	6	13	4
Race						
White	1	1	0	0	1	0
American Indian	1	0	0	0	1	0
Black	4	18	6	14	15	11
Black-Hispanic	0	0	0	0	0	0
Chinese	0	0	0	0	0	0
Filipino	0	0	0	0	1	1
Hawaiian	0	0	0	0	1	0
Hispanic	93	82	94	86	83	88
Other Asian	1	0	0	0	0	0
10TH GRADE						
English Language Status						
English Language Learner (ELL)	25	32	30	31	21	20
Not ELL	75	68	70	69	79	80
Gender						
Female	35	54	48	55	45	50
Male	65	46	52	45	55	50
Free/Reduced-Price Lunch (FRPL)						
FRPL Eligible	68	84	99	86	86	94
Not FRPL Eligible	32	16	1	14	14	6
Race						
White	0	0	1	0	0	0
American Indian	1	0	0	0	0	0
Black	1	17	8	17	33	15
Black-Hispanic	0	0	1	1	1	0
Chinese	1	0	0	0	0	0
Filipino	2%	0%	0%	0%	0%	0%
Hawaiian	0%	0%	0%	0%	0%	0%
Hispanic	94%	83%	91%	82%	66%	85%
Other Asian	0%	0%	0%	0%	0	0%

Michael & Susan Dell FOUNDATION

TABLE A2. TEST OF BASELINE EQUIVALENCE: STUDENTS IN THE ANALYTIC SAMPLE

PRETEST SCORES	BLAST			COMPARISON			BASELINE DIFFERENCE (BLAST - COMPARISON)	
	N	MEAN	SD	N	MEAN	SD	ES	P
CST MATHEMATICS								
Grade 9 (2011 Grade 8 CST Algebra I)	212	288.6	41.6	313	284.7	39.0	0.10	0.689
CST ELA & ELA BENCHMARK								
Grade 9 (2011 Grade 8 CST ELA)	316	325.2	54.8	405	319.6	48.3	0.13	0.394
Grade 10 (2010 Grade 8 CST ELA)	173	311.2	48.3	264	315.1	48.8	-0.08	0.717

Note: Effect size (ES) = Difference in means/pooled standard deviation (SD) between BLAST and comparison schools.

Teacher Survey Results - Alliance College-Ready High Schools

SURVEY ADMINISTRATION

Overall, 22 of 26 Alliance teachers implementing blended learning at the three BLAST schools completed the Teacher Survey for an 85% response rate. Table S1 describes the characteristics of the respondents in terms of the subject area(s) they teach and years of teaching experience. Additionally, 23 of 35 teachers at non-BLAST schools completed the Comparison Teacher Survey for a 66% response rate.

TABLE S1. TEACHER SURVEY RESPONDENT CHARACTERISTICS

TEACHER CHARACTERISTICS	NUMBER	PERCENT
SUBJECT TAUGHT		
English language arts (ELA)	5	23
Math	6	27
Science, social students, foreign langauge	11	50
TEACHING EXPERIENCE		
New teachers (3 years or less)	15	68
Veteran teachers (4 or more years)	7	32

Classroom Instructional Activities

ENGLISH LANGUAGE ARTS

English language arts (ELA) teachers implementing the BLAST model reported devoting between 16–30 minutes of a typical 2-hour ELA classroom session to teacher-led whole-class instruction (80%), teacher-led small-group instruction (100%), and small-group collaborative projects (80%), with lower numbers of teachers (20%) devoting the same amount of time to teacher-led one-on-one instruction. These ELA teachers also reported devoting over 30 minutes of a typical ELA classroom session to independent student work and small-group collaborative projects (20% for both) (see Figure S1).

MATH

Math teachers implementing the BLAST model reported devoting between 16–30 minutes of a typical 2-hour math classroom session to independent student work (83%), teacher-led small group instruction (67%), and small-group collaborative projects (67%), with smaller numbers of teachers devoting the same amount of time to teacher-led whole-class instruction (50%), teacher-led one-on-one instruction (33%), and student-directed instructional activities (50%). These math teachers also reported devoting over 30 minutes of a typical math classroom session to student-directed instructional activities (50%), small-group

FIGURE S1. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (ELA CLASSROOMS)

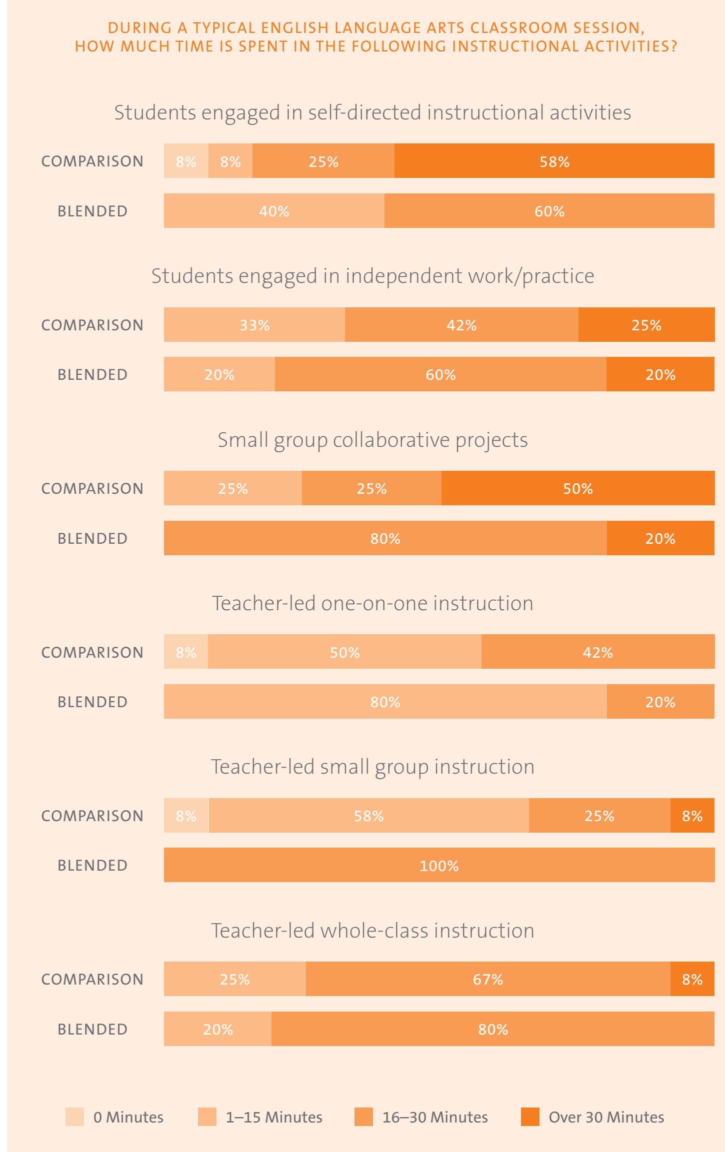


FIGURE S2. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (MATH CLASSROOMS)

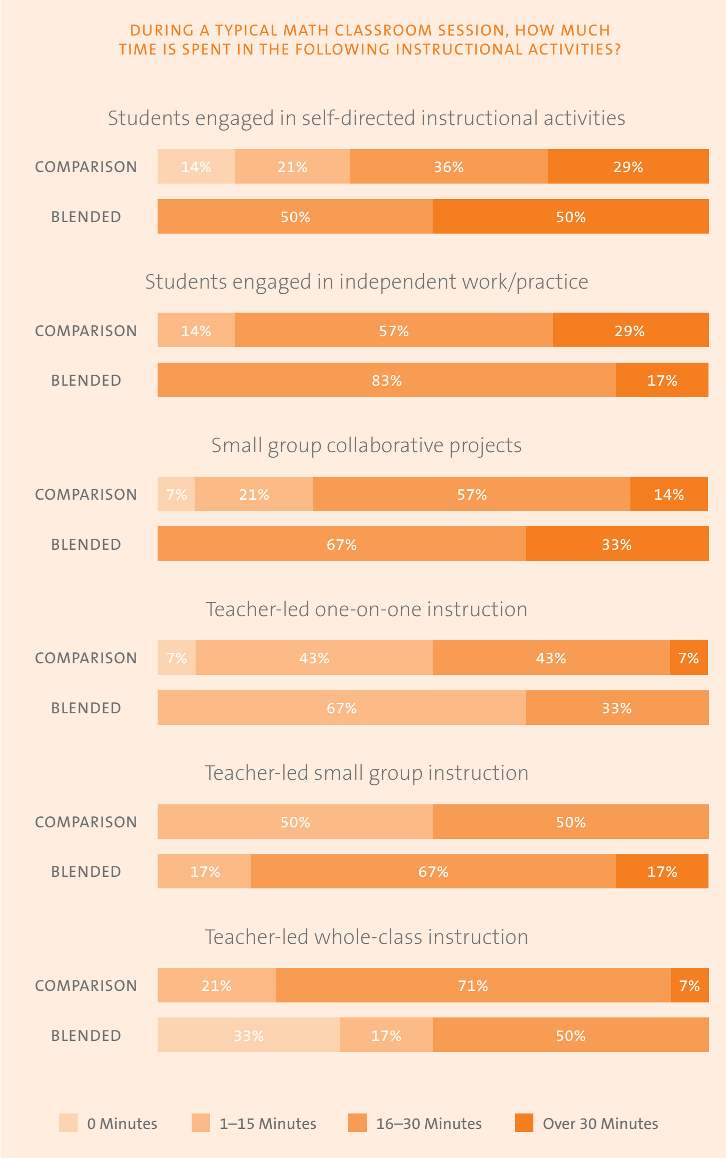
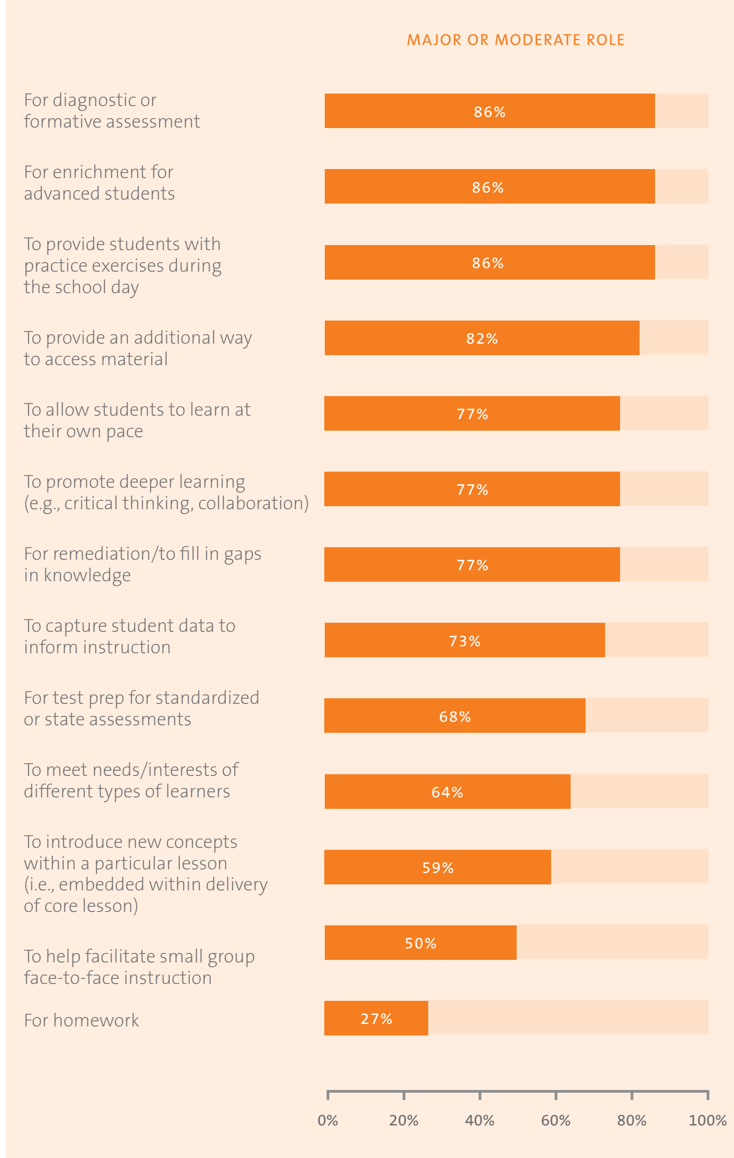


FIGURE S3. PERCENT OF TEACHERS REPORTING DIFFERENT ROLES FOR TECHNOLOGY AND WEB-BASED INSTRUCTION IN SUPPORTING TEACHER-PROVIDED INSTRUCTION



collaborative projects (33%), and teacher-led small-group instruction and independent student work (17% for both) (see Figure S2).

ELA and math teachers in the BLAST schools spent more time providing opportunities for students to engage in small-group collaborative projects and instruction compared to teachers in non-BLAST schools. However, these same BLAST teachers reported that they provide fewer opportunities, compared to non-BLAST teachers, for students to engage in self-directed learning activities and, in ELA classrooms, less time working with students one-on-one compared to their non-BLAST peers.

ROLE OF TECHNOLOGY

Teachers implementing the BLAST model reported that technology and web-based instruction played a major or moderate role in supporting instruction by "personalizing" the learning experience (e.g., providing an additional way to access material, allowing students to learn at their own pace, providing enrichment/remediation for advanced/struggling students), providing opportunities for promoting deeper learning, and helping capture and use data on student performance and achievement (see Figure S3).

Technology played a similar role across subject areas, with the exception of capturing data to inform instruction playing relatively more of a role in math (100% vs. 60% in ELA and 64% in math), and facilitating small-group instruction playing relatively more of a role in science, history, and foreign language (73% vs. 20% in ELA and 33% in math). Technology played far less of a role in ELA in introducing new concepts or meeting the needs/interests of different types of learners compared to the remaining subject areas (40% vs. 67% math and 64% in other subjects).

Both teachers implementing the BLAST model and those in the comparison group used technology and web-based instruction to "personalize" the learning experience, albeit in slightly different ways. For teachers implementing the BLAST model, technology and web-based instruction was used to provide additional ways to access material, allow students to learn at their own pace, and provide enrichment or remediation for advanced or struggling students. For teachers in the comparison group, technology and web-based instruction was used to provide additional ways to access material and remediation for struggling students and, more generally, to help meet the needs and interests of different types of learners. Both groups of teachers also reported using technology and web-based instruction to capture student data to inform instruction, with a higher proportion of BLAST teachers than comparison teachers reporting this type of use. One of the biggest differences between teachers implementing the BLAST model and those in the comparison group was in the use of technology and web-based instruction to help provide more opportunities for deeper learning—over 75% of BLAST teachers reported using technology to support deeper learning, whereas only 25% of comparison teachers reported the same type of use.

USE OF DATA

Almost three-quarters (73%) of the teachers implementing the BLAST model reported that they reviewed student performance data provided by the software system(s), although there is a disparity in how frequently teachers reviewed student reports. Among teachers who did not review the data, or reviewed it once a month or less, the most cited reasons included not having access to such a program (67%), the data not being informative or relevant (33%), or relying more on information outside the system (33%). In particular, Alliance teachers teaching subjects other than math and ELA who did not review the data indicated that they did not have access to a program in their subject area that collects and/or displays such data.

Of those who reviewed the data, half looked at the data two to three times a week or more frequently (50%); the other half looked at the data two to three times per month or more frequently (50%) (see Table S2). Alliance math teachers reported reviewing the data on a somewhat more frequent basis than teachers of other subjects, with 50% indicating they do so on almost a daily basis.

TABLE S2. FREQUENCY OF DATA REVIEW BY ALLIANCE TEACHERS

HOW OFTEN DO YOU LOOK AT THE DATA RECORDED BY THE SYSTEM(S) USED TO SUPPORT BLENDED LEARNING?	ELA (%)	MATH (%)	OTHER (%)	OVERALL (%)
Every day or almost every day	0	50	33	**31**
Two to three times a week	25	0	33	**19**
About once a week	75	33	33	**44**
Two to three times a month	0	17	0	**6**
About once a month or less	0	0	0	**0**

Seventy percent found the student performance data somewhat or very useful for informing their instruction, while just under one-third reported finding the student performance data only slightly useful. Alliance math teachers reported the student performance data captured by the software program(s) was more useful for them than for other teachers, particularly those teaching science, social studies, or foreign language.

Teachers implementing the BLAST model reported that reviewing student performance data from program(s) led them to monitor, diagnose, and provide feedback on individuals' and groups' understanding of key concepts, and modify plans for future lessons or instructional activities (see Figure S4). Teachers' uses of data were reasonably similar across subject areas. Consistent with science, social students, and foreign language teachers' use of technology to facilitate small-group instruction, a higher percent indicated they used data from the programs to form these groups (71% vs. 50% in math and ELA).

Michael & Susan Dell FOUNDATION

FIGURE S4. PERCENT OF TEACHERS USING DATA "SOMEWHAT" OR "A GREAT DEAL" FOR DIFFERENT INSTRUCTIONAL PURPOSES

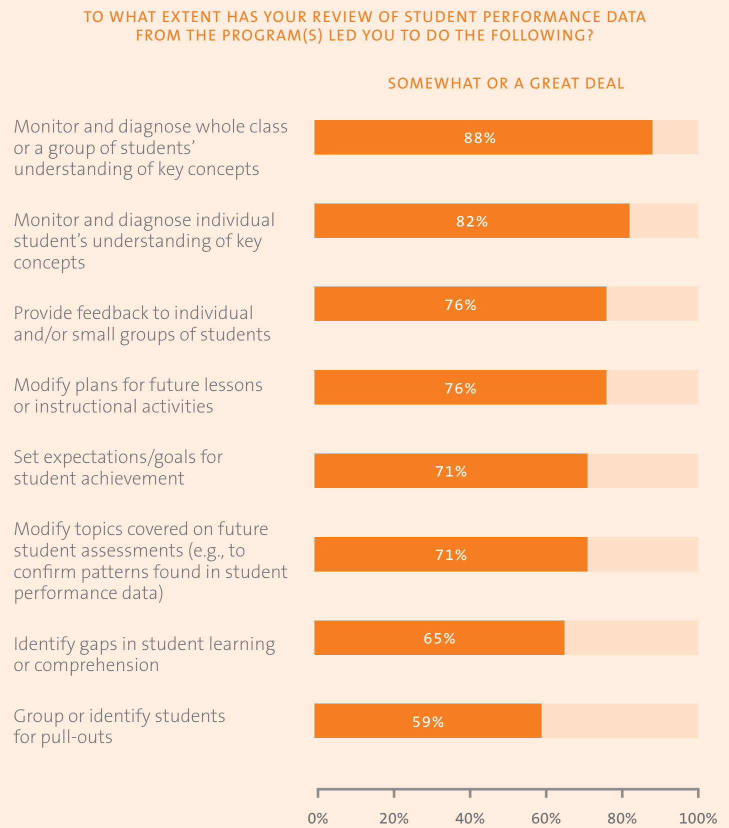
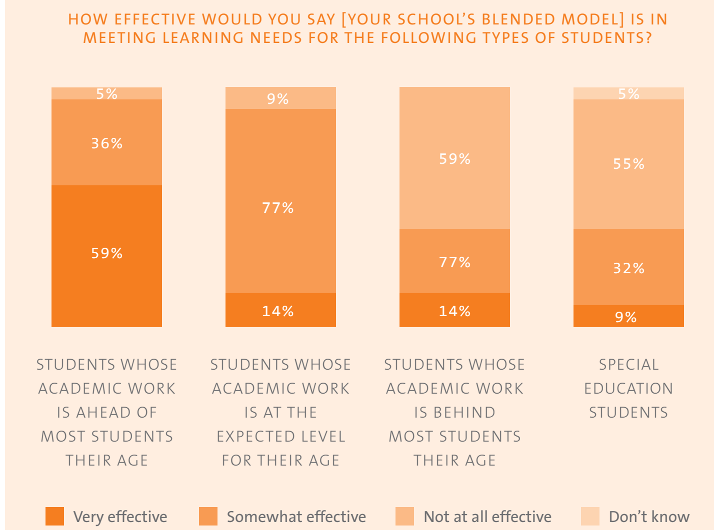


TABLE S3: PERCENTAGE OF TEACHERS REPORTING THEY "AGREE" OR "STRONGLY AGREE" THAT BLENDED LEARNING BENEFITS THEIR STUDENTS IN DIFFERENT WAYS.

PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS.	PERCENT AGREE OR STRONGLY AGREE ELA (%)	MATH (%)	OTHER (%)	OVERALL (%)
The BLAST model meets the learning needs of my students.	20	33	55	**41**
Students are highly engaged while using the BLAST model.	60	50	45	**50**
The BLAST model helps students take ownership for their own learning.	60	83	63	**68**
Students' learning and understanding of the materials has improved due to the use of the BLAST model.	20	50	45	**41**
I would recommend the use of the BLAST model to other teachers.	40	50	45	**45**

FIGURE S5. PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF EFFECTIVENESS OF THE BLENDED LEARNING MODEL FOR STUDENTS WITH DIFFERENT LEVELS OF ACADEMIC PERFORMANCE OR NEEDS.



Satisfaction with Blended Learning and Impact on Student Learning

TEACHER SATISFACTION WITH BLENDED LEARNING

Whereas two-thirds (68%) of teachers implementing the BLAST model agreed that the model helped students take ownership over their learning, half or less of the BLAST teachers agreed that students were highly engaged while using BLAST, that BLAST helped improve student learning, and that it met the learning needs of their students (see Table S3).

ELA teachers were the least positive in their responses regarding the BLAST model improving students' learning and understanding. The majority of math teachers (83%) agreed the BLAST model helped students take ownership for their own learning; only two-thirds of teachers in other subject areas felt the same way.

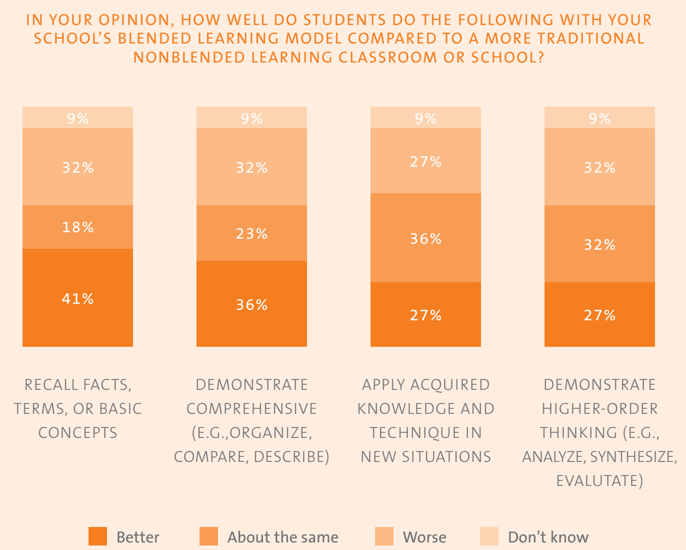
EFFECTIVENESS FOR STUDENT TYPES

While a majority of teachers implementing the BLAST model reported that the model was very effective for students whose academic work was ahead of most students their age, and somewhat effective for students whose academic work was at the expected level for their age, a majority also reported that the model as not at all effective for students whose academic work was behind most students their age, or for special education students (see Figure S5).

EFFECTIVENESS RELATIVE TO NONBLENDED ENVIRONMENT

In comparison to the skills acquired by students prior to the adoption of the BLAST model, about a third of BLAST teachers reported that students were worse off across a variety of skills, from the basic recall of facts to higher-order thinking (see Figure S6).

FIGURE S6. PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF THE EXTENT TO WHICH THEIR STUDENTS DEMONSTRATE HIGH-ORDER THINKING AS A RESULT OF THE BLENDED LEARNING MODEL RELATIVE TO A MORE CONVENTIONAL INSTRUCTIONAL MODEL.



Veteran teachers (4 or more years of teaching experience) at Alliance were more polarized in their ratings of students' skills under blended learning as compared to a more traditional model; they were often split down the middle with nearly half saying students did better in these areas while the other half that students did worse with blended learning. Compared to the veteran teachers, more of the new teachers (3 or fewer years of teaching experience) said students did about the same with or without blended learning, or that they didn't know.

Additionally, math teachers at Alliance were consistently more positive about the impact blended learning had on students' skills in a variety of areas, with half or more indicating students did better in these areas with blended learning than in a more conventional, nonblended environment. ELA teachers often indicated students did about the same with blended learning, or even worse with blended when it comes to recalling facts, terms, or basic concepts.

Training and Support

The majority (91%) of Alliance teachers participated in a training or orientation session (either in person or online) directly related to their school's use of blended learning or on the specific software program(s) supporting it. Of those who participated in a training or orientation session, 40% reported being very satisfied or satisfied with the training while 60% reported being dissatisfied or very dissatisfied with the training. Teachers cited training that was not specific to their content area and/or training that was too theoretical and did not include enough concrete examples or connections to the classroom as reasons for their dissatisfaction.

Almost 80% of teachers implementing the BLAST model reported spending between 1–20 hours of their own time (i.e., outside of professional development) getting acquainted with the software program(s) supporting the model, or planning for how to best integrate the model with their instruction. However, just over 20% of teachers implementing the BLAST model reported spending over 20 hours of their own time on this task (see Table S4).

TABLE S4. TIME SPENT OUTSIDE OF PROFESSIONAL DEVELOPMENT PREPARING FOR IMPLEMENTATION.

HOW MUCH OF YOUR OWN TIME (I.E., OUTSIDE OF ANY PROFESSIONAL DEVELOPMENT) DID YOU SPEND GETTING ACQUAINTED WITH THE SOFTWARE PROGRAM(S) SUPPORTING THE BLAST MODEL OR PLANNING FOR HOW TO BEST INTEGRATE THE BLAST MODEL WITH YOUR INSTRUCTION?	PERCENT
Less than 1 hour	9
1–5 hours	18
6–10 hours	32
11–20 hours	18
21–30 hours	5
Over 30 hours	18

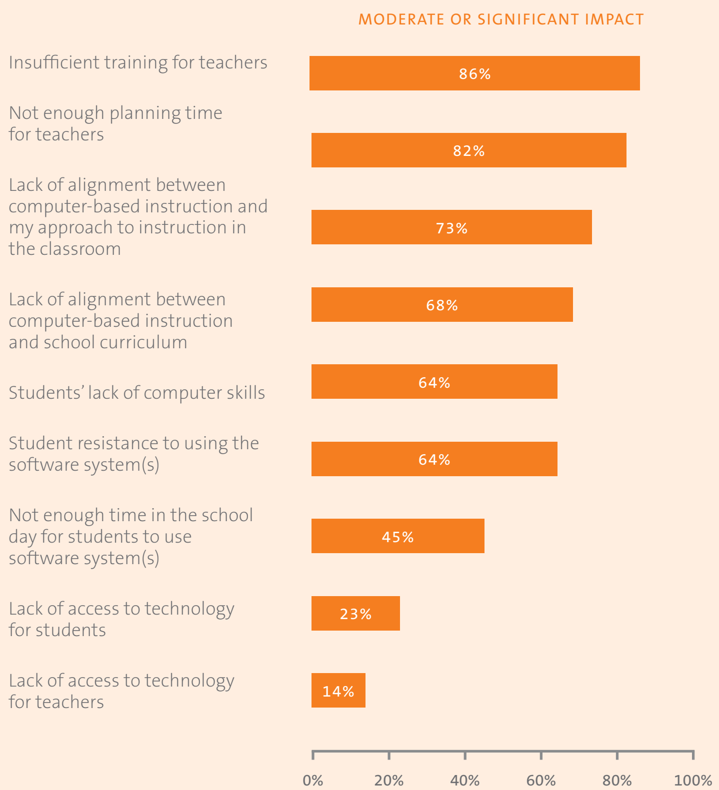
Math teachers at Alliance were the most positive in terms of their level of satisfaction with the training they received, while teachers in science, social studies, and foreign language were the least positive. Teachers in these other subject areas also reported spending more of their own personal time outside of professional development preparing to implement the BLAST model.

FACTORS INFLUENCING USE

Over two-thirds of teachers implementing the BLAST model reported that insufficient training, not enough planning time, and the lack of alignment between computer-based, and personal- and curriculum-guided approaches to instruction all had moderate or significant impacts on their, or their school's, ability to effectively use blended learning with their students (see Figure S7).

FIGURE S7. PERCENTAGE OF TEACHERS REPORTING THE EXTENT TO WHICH DIFFERENT FACTORS INFLUENCED THE EFFECTIVE USE OF THE BLENDED LEARNING MODEL TO A "MODERATE" OR "SIGNIFICANT" DEGREE.

PLEASE INDICATE THE DEGREE OF IMPACT THE FOLLOWING HAVE HAD ON YOU OR YOUR SCHOOL'S ABILITY TO EFFECTIVELY USE BLENDED LEARNING WITH YOUR STUDENTS.



Science, social studies, and foreign language teachers reported more challenges than ELA and math teachers related to planning time and alignment between the computer-based instruction and their approach to instruction in the classroom. ELA teachers reported the least challenge associated with student resistance to using the software program(s) (see Table S5). Furthermore, a lower percent of ELA teachers cited technical challenges as a moderate or significant barrier than teachers in the remaining subject areas (20% versus over 50%, respectively).

TABLE S5. FACTORS INFLUENCING EFFECTIVE USE BY SUBJECT AREA.

PLEASE INDICATE THE DEGREE OF IMPACT THE FOLLOWING HAVE HAD ON YOUR ABILITY AND/OR YOUR SCHOOL'S ABILITY TO EFFECTIVELY USE BLENDED LEARNING.	PERCENT MODERATE OR SIGNIFICANT IMPACT		
	ELA (%)	MATH (%)	OVERALL (%)
Insufficient training for teachers	80	83	91
Not enough planning time for teachers	60	67	100
Lack of alignment between computer-based instruction and the school curriculum	60	67	73
Lack of alignment between computer-based instruction and my approach to instruction in the classroom	60	50	91
Student resistance to using the software system(s)	40	66	73

Nearly all (95%) of teachers reported that they or their students experienced technical problems in using the BLAST model. Of those teachers reporting technical problems, half (52%) reported that the technical problems were only a slight barrier to their effective use of the BLAST model, while the other half (48%) reported that the technical problems were either a moderate or significant barrier to their effective use of the BLAST model. The most commonly cited technical problems included student computers not working properly (86%), Internet connection unreliable/down (71%), software glitches or bugs (67%), and incompatibilities with existing firewalls or filters set up by district or school (57%).

Michael & Susan Dell FOUNDATION

Research Profile: FirstLine Schools
Arthur Ashe Charter School (2011–12 School Year)

Introduction

Since its inception in 2007 as a charter management organization (CMO), FirstLine School's mission has been to create and inspire open admissions public schools in New Orleans that prepare students for college and fulfilling careers by focusing on college-readiness, providing a rich variety of experiences for students, and developing skillful staff and sustainable organizations (www.firstlineschools.org). There are currently four elementary schools in the FirstLine network—Arthur Ashe (Ashe), Langston Hughes Academy, Samuel J. Green Charter School, and John Dibert Community School—and one high school, Joseph S. Clark Preparatory High School (Clark), all serving students in New Orleans.

FirstLine's motivation for piloting a blended learning model in two of its schools in 2011–12—Ashe and Clark—was to provide more personalized academic support for students through a combination of adaptive online instruction and small-group teacher-led instruction. This report focuses on the experience of teachers and students within Arthur Ashe Elementary School. Online-instruction for a majority of students took place within a learning lab, which was used to facilitate one-on-one and small-group instruction by giving teachers more time to work with students with the greatest academic needs while the other students were engaged with various online instructional programs. In this way, FirstLine used blended learning to support the implementation of both its Response to Intervention (RTI) program and instruction for students with special learning needs.[15] The online instruction element of the model was meant to supplement the core teacher-led classroom instruction, and there was very little, if any, formal integration with the curriculum. Instead, the online instruction was used to "fill in the gaps" and provide extension material for accelerated students while the primary instruction came from the teacher.[16]

The 2011–12 school year was the first year the blended learning model was implemented in Arthur Ashe. Ashe was opened in 2007 and originally served students in Grades 3–6 before expanding in 2009–10 to Grades K–8. For the 2011–12 school year, Ashe experienced a significant change in enrollment. Approximately 100 new students enrolled in the school in the fall, while about 25% of the students in the prior cohort left the schools. (More than half of the 112 public schools in New Orleans are "open enrollment" schools, which encourages student mobility.) Special education students accounted for over 26% of Ashe's enrollments, the highest percentage in the city, and almost 40% of Ashe's students had Individualized Educational Plans (IEPs). Nearly all (97%) of Ashe's students were eligible for the federal free or reduced-price lunch program.

BLENDED LEARNING MODEL DESCRIPTION

During the 2011–12 school year, Ashe implemented a lab rotation blended learning model. There were two designated labs in the school, each with 60 student computers, able to support two classes at a time. In Grades K–3, students spent 30 minutes per day in an English language arts (ELA) lab and 30–40 minutes per day in math lab; for Grades 4–8, students spent 50 minutes every other day in the ELA lab and 50 minutes per day in math lab.[17] In addition to time in the learning labs, students spent about 5 hours per week receiving teacher-led instruction in both ELA and math in the classroom. According to the principal, this represented a decrease in teacher-led classroom instruction of about 30% compared to the 2010–11 school year, the year prior to the adoption of blended learning at Ashe (from 100 minutes per day to 70 minutes per day). Students' use of the different online instructional programs in the lab varied by subject area, grade level, and academic needs.

The schools' use of blended learning to facilitate one-on-one and small-group instruction was a key component of Ashe's RTI and special education program. During lab time, RTI Tier 2 students might receive additional teacher-led instruction, either within the lab ("push in") or in a nearby classroom ("pull-out"). During these sessions students might work on laptops on different software programs than the other students in the learning lab, including SuccessMaker, Lexia, ReadingPlus, Reading Eggs, and DreamBox (SuccessMaker is the primary online program used by Tier 2 students). Students with the greatest academic needs at Ashe—RTI Tier 3 and special education students—spent their lab time in separate "learning support" classrooms staffed with a core subject area teacher and a special education specialist. These students typically spent their lab time rotating between three different stations: small-group, teacher-led instruction; online instruction; and independent practice. VmathLive and Ticket to Read were the primary online programs used in the learning support classroom.

Each lab had a designated lab coach or monitor. These individuals were paraprofessionals, rather than certified teachers, and largely responsible for classroom management and establishing an effective learning culture with clear norms and expectations for student behavior and engagement. In addition, the lab coach was responsible for the academic progress of the Tier 1 students while in the lab. They held parent-teacher conferences and provided the pass/fail grade for Tier 1 students' work in the lab. Student-to-student communication in the learning lab was limited and generally discouraged. During lab time, one or two subject-specialty core teachers and special education teachers were also present to supervise and support the school's RTI program. These teachers spent their time in the lab supporting students in different ways. They provided academic support to individual students who might be having difficulty with some aspect of the online content or provided direct remedial

---

[15] *Response-to-Intervention programs are a set of practices to help schools continuously identify students with different learning needs and provide interventions for students with the greatest needs. Schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions, and adjust the intensity and nature of those interventions depending on a student's responsiveness. Students at Ashe were grouped into three tiers based on learning needs and different types of instruction and intervention were provided to each. Tier 1 students were at or above grade level and did not require intervention; Tier 2 students were 1–3 years behind their peers; and Tier 3 students were 3 or more years behind their peers. In 2011–12 approximately 28% of Ashe's students were designated as Tier 1, 53% Tier 2, and 19% Tier 3.*

[16] *A detailed description of FirstLine's instructional and operational model is available in a case study developed as part of this research study and published on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).*

[17] *By the end of the year, the lab time for K–3 students was combined into a single 60-minute block—split between ELA and math—to reduce the number of daily transitions between classrooms for students.*

TABLE 1: FIRSTLINE SCHOOLS IN RESEARCH SAMPLE				
SCHOOL NAME (YEAR OPENED AS A FIRSTLINE SCHOOL)	GRADE LEVELS	STUDENTS (#)	STATUS OF BLENDED LEARNING SY 2011–12	2011–12 STATE REPORT CARD GRADE
Arthur Ashe Charter School (2007)	K–8	460	Full Year	D
John Dibert Community School (2010)	K–8	430	Partial Year	T (Top Gain)
Langston Hughes Academy (2010)	K–8	630	Partial Year	D+
Samuel J. Green Charter School (2006)	K–8	530	No	D

TABLE 2: DATA COLLECTION SUMMARY, FALL 2011 AND SPRING 2012 SITE VISITS, ARTHUR ASHE		
	FALL 2011 SITE VISIT	SPRING 2012 SITE VISITS
OBSERVATIONS		
Classrooms	1	4
Learning Labs	3	2
INTERVIEWS		
CMO[A] Leaders	1	2
School Leaders	1	1
Teachers	1	4
Lab Coaches	2	1
Students	0	8

TABLE 3: RESPONSE RATES, TEACHER SURVEY, FIRSTLINE ASHE	
INSTRUMENT	RESPONSE RATE
Ashe Teacher Survey	95% (19 of 20)
Lab Coach Survey	100% (2 of 2)
Comparison Teacher Survey	65% (47 of 72)

[A] *Charter Management Organization*

instruction to an individual or small group of students in a section of the lab set aside for tutoring or in a nearby classroom.

To help motivate students, Ashe displayed the progress reports on the walls of the learning lab. They also promoted healthy competitions among individual students and across classes, awarded prizes for the most progress, and took away privileges (e.g., missing special events on Fridays) for those students who failed to make adequate progress. To motivate the middle school students, teachers encouraged goal tracking and students received pass/fail grades for their work in the lab.

Sample

FirstLine Schools currently has four charter elementary schools in its network all located in the New Orleans Recovery School District. Arthur Ashe Elementary Schools adopted a blended learning model in the school year (SY) 2011–12. Two of the other three schools acting as comparisons—John Dibert Community School and Langston Hughes Academy—experimented with some form of blended learning later in the school year, beginning as early as October for Dibert and November for Langston Hughes. Both schools implemented ST Math, with Dibert using a classroom-rotation model and Langston Hughes a rotation learning lab model. In both schools the focus of ST Math use was in the upper grade levels, Grades 4–8. Langston Hughes also implemented iStation into their K–4 ELA curriculum, beginning as early as November, as a regular in-class rotation.

All schools serve students from predominantly low-income communities within New Orleans, and all have open enrollment policies and are attended by students across the Recovery School District (New Orleans). Table 1 shows the size of the schools along with their 2011–12 performance "grade" from their School Report Card.[18] Arthur Ashe performance scores were in the bottom third of schools in the state as were two of the comparison schools. John

[18] *Louisiana adopted letter grades based on School Performance Scores in 2010. All schools with sufficient data receive school performance scores. For elementary and middle schools, 95%–100% of the School Performance Score is based on state assessments that are administered each spring. Letter grades range from A to F, and F schools are considered failing (http://www.louisianabelieves.com/accountability/school-letter-grades).*

Michael & Susan Dell FOUNDATION

TABLE 4: CORE ONLINE PROGRAMS FOR ELA AND MATH, ASHE ELEMENTARY SCHOOL

ELA PROGRAMS

GRADE LEVEL	SUCCESSMAKER	iSTATION	ACHIEVE3000	
K–1		X		
2–3	X	X		
4–8	X	X	X	

MATH PROGRAMS

GRADE LEVEL	SUCCESSMAKER	DREAMBOX	ST MATH	APANGEA
K–1		X	X	
2–3	X	X	X	
4–5	X		X	
6–8	X		X	X

Dibert Community School was designated a Top Gain school (T) because it exceeded its achievement growth target.[19]

Data Collection Summary

SITE VISITS

Members of the research team conducted a site visit to Arthur Ashe in October 2011 and May 2012. Site visit activities are summarized in Table 2.

TEACHER SURVEY

During spring 2012, the research team conducted surveys of teachers and learning lab facilitators in Ashe. Survey participation rates are summarized in Table 3.

SYSTEM LOG DATA

System log data were provided by FirstLine for students enrolled in Arthur Ashe in SY 2011–12. For the purpose of summarizing these data we aggregated the log data across the following grade levels: K–1, 2–3, 4–5, and 6–8. For an analysis of log data being used to support math instruction, system data were accessed from ST Math, DreamBox, SuccessMaker, and Apangea. For ELA instruction, we accessed and analyzed log data from iStation, Achieve3000, and SuccessMaker. Table 4 shows how the most frequently used online programs for ELA and math were distributed across grade levels and subject areas.

For the purpose of these analyses, only students' interactions with the most frequently used online instructional programs for ELA and math were analyzed. Use data was not accessed or analyzed from other digital instructional resources used less frequently in the learning lab (e.g., Reading Eggs, ReadingPlus, Grammar Fitness, Study Island, Lexia, VmathLive).[20] Thus students' use of online instruction programs reported here might not represent the full extent of technology-supported instruction a student received in the learning lab. We also did not capture students' use of online resources outside these programs that might be used in the classroom, including tools to conduct online research, collaboration tools, and applications used to produce learning products (word processing, multimedia production) and analyze data.

[19] *To recognize schools that make significant growth from one year to the next, Louisiana designates schools that achieve their growth target as a Top Gains school if they are not in subgroup component failure, and they are not identified as failing No Child Left Behind (NCLB) subgroup performance more than one year. Top Gains schools are eligible for monetary awards that can be used for any educational purpose within the school (http://www.louisianabelieves.com/accountability/top-gains-schools).*

[20] *VmathLive was the primary math program used by Tier 3 students in the learning-support classroom.*

STUDENT LEARNING OUTCOMES

Due to the high level of missing data for some student outcome measures for Grades 4–8 within Ashe, the analyses of impacts on student learning was limited to Grades 1–3 and scores on the math and ELA Terra Nova assessments. We include in the analyses of student test scores only students that have both spring 2011 and 2012 scores. The spring 2011 score is used in the analysis models to control for existing differences in prior achievement between the students attending Ashe and the comparison schools. If more than 25% of a grade-level cohort was missing an outcome score or a measure of prior achievement (1 in 4 students), the grade level was dropped from the analysis.[21]

Implementation Findings

Findings from the site visits and teacher survey were used to help understand the impact of the use of the learning lab on classroom instruction relative to other schools in the FirstLine network, as well as the facilitating factors and barriers to implementing the school's model during the 2011–12 school year. Additionally, the student log data from the online programs was used to help better understand the intensity with which students used the primary online programs supporting the blended learning model for math and ELA instruction.

FINDINGS FROM THE SITE VISITS AND TEACHER SURVEY

Data from the fall 2011 and spring 2012 site visits and the teacher survey were used to report on different aspects of implementation of FirstLine's learning lab rotation model, including differences in classroom instruction between blended and nonblended schools, infrastructure and technology issues, the support and training teachers received and their role in the learning labs, characteristics and quality of system software, and teachers' use of the system data to support their instruction. In addition we report on teachers' overall satisfaction with the learning lab model and the perceived benefits to students and teaching based on teacher reports and through the research team's own observations and interviews with administrators, teachers, and lab coaches.[22] A summary of the survey results is included in the Appendix for this profile.

Infrastructure and Technology

Students had one-to-one access to computers in and outside the learning labs. The learning labs had enough computers available to support one-to-one computing for each student, such that each student could work on a computer independently with a wireless connection to the Internet. The labs could accommodate up to 60 students in a learning lab at one time. There were also laptops available so that students participating in "pull-out" sessions with a teacher during lab time could use the computer at the teacher's discretion (pull-outs could contain a mix of computer-mediated and face-to-face instruction). In the learning support classrooms (RTI Tier 3 and special education students) there were five to six computers that students used in a rotation model, moving between online instruction, small-group instruction with the teacher, and working independently on practice worksheets. Some ELA classrooms also had five to six computers that supported "literacy centers" and provided students access to e-books and online literacy programs through providers like Raz-Kids (http://www.raz-kids.com).[23]

Issues with connectivity and bandwidth impacted use of blended learning. All teachers completing the survey reported that they or their students experienced some technical problems during the school year with the most prevalent issues being unreliable Internet connectivity, insufficient bandwidth, broken computers, and unreliable software. Almost half the teachers indicated that technical challenges posed a moderate or significant barrier to their effective use of the blending learning model. FirstLine leadership

---

[21] *We intended to include in our analysis of impacts on learning outcomes scores on the Louisiana state achievement tests as well as assessments developed by Louisiana's Achievement Network (a nonprofit organization that works with schools in low-income communities to develop a culture of using learning standards and achievement data to improve academic achievement). However, in both cases the percent of students attending Ashe with both spring 2011 and 2012 scores were well below the 75% inclusion criteria for all grade levels.*

[22] *The lab coaches were also surveyed, but the sample size is too small to report. Therefore, the survey data reported represents the teachers' perspective, though in at least some cases learning coaches were the primary actors in implementing the online component of the model.*

[23] *According to school and FirstLine leadership, during the 2012–13 school year, with the school in its new building, the goal was to place five to six computers in each classroom to allow teachers to bring blended learning into the classroom. In addition, FirstLine planned to move ELA instruction out of the learning lab entirely. Instead, students will have access to computers within the library and will rotate through a "literacy center" supported with online instruction focusing on grammar, vocabulary, and mechanics; an independent reading station using print-based books from the library; and another station for independent work. A teacher will be present to provide individual and small-group academic support, while a lab coach will be responsible for maintaining a positive learning culture.*

acknowledged that there were issues throughout the year with Internet bandwidth—some residing with the school and others with the software providers—that impacted students' use of the online program. FirstLine's Director of Blended Learning (DBL) estimated that Ashe lost about 3–4 weeks of instruction in the learning lab across the year due to intermittent loss of the Internet or software problems with individual online programs.

On-site IT support and back-up instructional activities were needed to anticipate problems with technology. Teachers and administrators indicated the importance of on-site IT support familiar with the software programs. Ashe had a TEACH-UP Intern who provided day-to-day, on-site IT support, and was often present when the learning labs were in use. FirstLine also employed an additional IT staff member who supported the entire FirstLine school network and was available to schools when larger technical issues arose. The DBL also expressed the importance of having access to multiple online programs and offline activities available to students in the learning lab in the event the Internet went down or there were issues with the performance of a particular online program.

Desktop monitoring software was used to limit off-task activities. To increase student's productivity in the learning lab, lab coaches reported using desktop monitoring software to limit students' off-task activities. In addition to regularly reviewing real-time images of students' desktops on her own computer, one lab coach projected her display on a screen at the front of the lab as a reminder to students that she was monitoring whether they were working on the programs they were assigned as well as the progress they were making.

The support provided by FirstLine's Director of Blended Learning was critical to the school's ability to effectively implement their blended learning model. The adoption of blended learning and the use of learning labs during the 2011–12 school year was a dramatic change for the school and teachers. School leaders noted that one of the most critical factors facilitating the implementation of blended learning in the school was the availability of a dedicated FirstLine staff member—the Director of Blended Learning (DBL)—to coordinate the roll-out of the model and thus reduce the burden on school leaders and, in particular, teachers. The DBL coordinated the piloting and evaluation of different software programs with teachers; negotiated purchases of hardware and software with vendors; coordinated professional development sessions provided by vendors and delivered customized trainings; compiled reports of student progress based on online system use data for teachers; and served as a liaison between technology and software vendors and the school. The DBL supported all schools in the FirstLine network that were piloting blended learning instruction.

SOFTWARE SELECTION, AVAILABILITY, AND DESIGN

FirstLine's product selection and evaluation process facilitated continuous refinement of its blended instructional model. FirstLine leadership stressed the importance of having clear criteria for the selection of software programs that helped decision-makers get beyond claims included in product marketing materials. Examples of such criteria included cost per student, bandwidth usage, content coverage and standards alignment, assignability, quality of reporting, and responsiveness and support from the vendor. FirstLine developed a rubric outlining required and desired characteristics and used it to make initial online program selections. Because the educational software available in the market was evolving quickly, FirstLine constantly reevaluated its own suite of online programs throughout the school year and continued to search the market for other alternatives that could better serve their students and teachers. Finally, FirstLine found it critical to the success of the roll-out of its blended learning model during the study year to involve teachers in the piloting and selection of all software; this allowed the teachers to provide feedback on the programs prior to purchase, as well as increased teacher ownership over the use of selected programs to support the model.

FirstLine leadership and school administrators all reported that they were more satisfied with the content coverage and quality of software programs to support math instruction than they were for those programs developed to support English Language Arts (ELA) instruction. Teachers reported that students found ELA programs that focus on reading comprehension (e.g., Achieve3000) much less engaging than the math programs they used, which tended to be more interactive. In programs like Achieve3000, students read selections of nonfiction text appropriate to their reading level on the computer monitor and then answered a series of reading comprehension questions. Within the learning lab, there was no opportunity for these students to discuss what they had read with peers or teachers.[24]

[24] *According to the DBL, the ELA instructional space in the school's new building will be designed to enable literacy discussions in small groups.*

The quality of the reporting on individual student progress within an online program varied significantly across programs, and in some, limited the school's ability to use specific programs to support the school's RTI program. According to the DBL, teachers working with RTI Tier 3 students needed to be able to closely monitor a student's progress against a list of specific skills that the student had yet to master. Few programs were available during the 2011–12 school year that provided this capability. SuccessMaker, VmathLive, and Ticket to Read were such programs and the primary programs used by FirstLine to support blended learning within the learning-support classroom. By the end of the school year one other program was added to this list, ST Math, which had refined its reporting of individual student progress in response to a request from the DBL.

To support instruction in the learning-support classrooms of students with special learning needs, software programs that allowed teachers to assign content was a critical feature. The assignability designed into some programs allowed teachers to customize the online-learning content for individual students based on their needs. An important factor in FirstLine's selection of SuccessMaker for its learning support classrooms was its assignability. ST Math is assignable to a limited degree; DreamBox is not assignable, and for this reason was not to be used in the school during the 2012–13 school year according to the DBL.

TEACHER TRAINING

Teacher satisfaction with the training provided on the school's blended learning initiative and online programs varied. Almost all teachers (95%) participated in a training or an orientation session (either in person or by webinar), in some cases provided directly by software vendors. Close to two-thirds (62%) were satisfied or very satisfied with the training they received, while more than a third (39%) were dissatisfied or very dissatisfied. Those who were not satisfied with the training they received generally indicated they needed more information about the software programs, particularly related to access to student data. ELA teachers were more satisfied than math teachers with the training they were provided. Math teachers in particular wanted more information about accessing and interpreting student performance data captured by the software programs.

Teachers were provided with regular opportunities throughout the school year to receive updates and provided administrators with feedback on the blended instructional model. Every Friday, teachers at Ashe had dedicated time for professional development from 1:30–4:00 p.m. According to the DBL, having this dedicated time each week set aside during which he could work with teachers, provide status updates, address their concerns, and receive their feedback on the overall blended learning model and individual programs was critical to the progress the school made in implementing the blended learning model during the 2011–12 school year. Although this time was heavily devoted to the implementation of blended learning towards the beginning of the school year (e.g., the DBL met with teachers and monitors weekly), towards the end of the year, the need to meet decreased to a monthly basis.

In addition to attending formal training, teachers also spent their own time getting familiar with the programs or planning for how to best integrate blended learning with their instruction, some more than others. Overall, about a third of the teachers spent 6 or more hours of their own time getting acquainted with the software programs, while two-thirds spent 5 hours or less. According to the DBL, in the future, FirstLine would like to distribute program log-ins to teachers during the summer so they can spend time exploring the new programs and coming up with ways to integrate the programs with their curriculum and instruction.

TEACHERS' ROLES IN THE LEARNING LAB

Teachers played a significant instructional role within the learning lab and used the time to provide direct one-on-one tutoring and small-group instruction. One or two subject-specialty core teachers, interventionists, and/or special education teachers were present in the lab to provide academic support to students working on the online programs when they were not providing remedial support to select students (teachers were often in the lab 2 days per week supervising and supporting students). Seventy percent of teachers reported they supervised students in the learning lab at least once a week (13 of 19). Eighty percent of teachers reported that they pulled out students from the lab for remedial one-on-one or small-group instruction at least twice a week, and almost half of the teachers reported doing so on a daily basis.

Although teachers were present in the learning lab, many teachers still felt that they didn't have a good sense of how students were performing and progressing within the online instructional programs. More than 80% of teachers reported that it was "very" or "fairly" important for teachers to know how students were performing in the learning lab to be a more effective teacher. Yet more than half the teachers (58%) reported they had little or no knowledge of what content areas students might be struggling with in the online programs with one in five teachers reporting they had no knowledge at all of how their students were progressing.[25]

During the 2011–12 school year, just over a third (37%) of teachers reported that they met with lab coaches once a week or more to discuss students' performance in the learning lab; about half (42%) did this three times a month or less, and a fifth (21%) never met with lab coaches. A majority of teachers (nearly 70%) reported they would have liked to meet with lab coaches more often than they currently met to discuss the progress and needs of individual students in their classrooms.

USE OF DATA TO INFORM INSTRUCTION

Teachers reported they regularly reviewed reports of student progress on the online programs and used the data to inform their instruction. Nearly all (95%) Ashe teachers reviewed the student performance data recorded by the software systems used to support blended learning, and did so once a week or more. FirstLine Ashe had dedicated "data days" every Wednesday that gave teachers a built-in opportunity to review system data from programs used to support RTI Tier 2 and Tier 3 students and determine appropriate next steps or interventions for their small-group instruction. As a result, Ashe teachers' review of the student performance data influenced their actions and decisions around instruction for individual students or small groups of students (e.g., setting goals for student achievement, grouping or identifying students for pull-outs, providing feedback to individual students) more so than instruction involving the entire class (e.g., modifying topics covered on future student assessments, modifying plans for future lessons or instructional activities). Three-quarters or more of the teachers indicated that data from the online reports played at least a moderate role in informing their decisions related to one-one-one or small-group instruction activities, compared to around half or fewer who used the data to inform whole-class actions.

Before the student progress data captured by the online programs could be leveraged for instructional purposes, several challenges had to be overcome to make the information more accessible to teachers. One challenge was related to the varying quality of the dashboards provided by the programs for monitoring student progress. According to the DBL, some programs—such as SuccessMaker—made it relatively easy for teachers to access and interpret student progress relative to specific content areas, but many did not. Another challenge facing teachers and lab coaches was how to make sense of the various output provided by the different systems since each system used a different metric and format for reporting student progress and performance. To support the school's use of data, the DBL estimated he spent about 2 hours each week aggregating the progress data across the different systems into a single report for teachers and lab coaches. In addition, he developed "report interpretation guides" for each of the programs so that school staff could more efficiently interpret the data reports. Without these efforts by the DBL, the usefulness of the system log data from the learning lab for diagnosing struggling learners and monitoring student progress would have been severely limited.

SATISFACTION WITH BLENDED LEARNING AND PERCEIVED IMPACTS ON STUDENTS

Teacher satisfaction with the blended learning model was high. Ashe teachers were highly satisfied with their first year experience with blended learning. Almost all teachers reported they were satisfied with the blending learning model at Ashe, with 95% agreeing they would recommend it to other teachers. Additionally, 90% of teachers believed the learning lab was meeting the learning needs of their students. In addition, half or more (47% to 63%) of the teachers reported that students were more adept at a variety of different skills ranging from basic recall of facts to higher-order thinking as a result of the school's blended learning program than in a more traditional, nonblended environment; few teachers (11% or less) reported that students' abilities in these areas were worse as a result of the adoption of the learning lab model.

Teachers perceived that all students benefited from the blended learning model, particularly students who worked above grade level. Most Ashe teachers believed the learning lab model was at least "somewhat" or "very effective" for almost all students (greater than 80%), particularly for students whose academic work was ahead of most students their age. Just over 60% of teachers reported that the blended model was "very effective" for students who were performing

[25] *By design, lab coaches were responsible for monitoring the progress made in the lab by Tier 1 students, those that were in the lab the bulk of the time. Core teachers were not expected to engage with the data for these students generated by the online programs in the lab. Core teachers were responsible for the progress of the Tier 2 students that they were pulling out of the learning lab for remediation. These teachers also had weekly Wednesday morning reviews of their students' progress in the Tier 2 programs. In general, core teachers were told by the DBL to not worry about the main content in the lab because that was the lab coaches' responsibility.*

above their grade level, compared to about 40% of teachers who reported the same for students at their grade level or below, including special education students.

However, there were differences between math and ELA teachers in the groups of students they thought benefited more from the blended learning model. A significant majority of math teachers indicated the school's blended learning model was "very effective" for students whose academic work was ahead of or at the expected level for their age (86% and 71%, respectively), a belief shared by half or fewer of the ELA teachers. In contrast, ELA teachers reported blended learning was more effective for students whose academic work was behind most students their age (56% reported the model was "very effective" for these students compared to 14% of math teachers).

Several nontechnology factors may be acting as barriers to effective use. Almost one-half of all teachers reported students' lack of computer skills, insufficient training for teachers, and lack of alignment between computer-based instruction and the school curriculum as the most prevalent nontechnology factors having a moderate or significant impact on the school's and teachers' ability to effectively implement the learning lab rotational model at Ashe. In addition, almost 4 in 10 teachers reported that students' resistance to using the software programs was also a moderate to significant factor hindering the effectiveness of the blended learning model.

FINDINGS FROM ANALYSIS OF SYSTEM LOG DATA

To capture the intensity with which students used the different core online learning programs, we computed the number of minutes that a student logged over the course of the school year for each program and grade level. As described above, the programs used differed by grade level and to some extent by prior academic performance. For example, in math in Grades 6 through 8, SuccessMaker was used more often by struggling math learners, while more advanced students were given more opportunities to use ST Math and Apangea. The classroom teachers made decisions about which programs students used based on the students' needs.

Tables 5–8 provide descriptive statistics of key use indicators, and Figures 1 and 2 illustrate how online instruction was used to personalize instruction within Ashe. The information presented in the tables include: the number of online programs used for ELA and math by students within the different grade levels (Table 5); the number of minutes spent on each core online program by grade level by subject (Tables 6 and 7); and the number of minutes spent by students by grade level by RTI level or tier (Table 8). In addition, Figures 1 and 2 show how the number of minutes that students spent on each online program varied by students' prior achievement scores (spring 2011 state achievement score).

A summary of the findings from the analysis of system log data follows.

Students in the upper elementary and middle school grade levels spent more time on the core online programs than younger students in Grades K–3. As shown in Tables 6 and 7 and Figures 1 and 2, the time spent in online instructional programs varies by grade levels. More time was spent in online instruction in Grades 4–8 compared to Grades K–3. Students in Grades 4–8 spent approximately 50% more time in online instruction than younger students in Grades K–3. In Grades K–1 and 2–3, the average time students spent in online instruction in ELA (Table 6) was 1,337 and 1,655 minutes, respectively (or about 40 to 45 minutes per week over a 36-week school year) compared to Grades 6–8 and 4–5, where students spent about 2,000 to 2,500 minutes, respectively (or 55 to 70 minutes per week). This trend is the same for math.

While the discrepancy between the amount of use of the core online programs between the younger and older students was expected for math (the daily learning lab sessions for math in Grades 4–8 were 50 minutes long compared to 30 minutes in Grade K–1) the discrepancy between grade levels for online ELA instruction was not expected. Both grade levels allocated approximately 150 minutes per week of learning lab time to ELA instruction each week. The discrepancy may have emerged from a combination of two different sources: older students may be more productive with their time in the learning lab than younger students (e.g., less off-task behavior, more efficient logging onto programs) and younger students may have been more likely than older students to work with online programs other than the core programs included in system log data analysis. While we collected no evidence during our observations or interviews with school administrators, teachers, or lab coaches that younger students tended to be less productive with their time in the learning lab than older students, we did learn from the DBL that many of younger students spent a significant amount of time on additional online programs other than the core programs, particularly at the beginning of the year. At the beginning of the school year, most students in K–2 were assigned to the early reading program Reading Eggs to build up basic reading skills before being allowed to "graduate" to iStation, the school's core online program for ELA. According to the DBL, many students were assigned to Reading Eggs for the first 4 to 6 weeks before being moved to iStation.

Students were assigned different combinations of online instructional programs depending on their academic needs. Students with different academic needs were assigned different combinations of online programs in an attempt to personalize their learning lab instructional experience. SuccessMaker was

TABLE 5: DESCRIPTIVE STATISTICS FOR NUMBER OF USERS, NUMBER OF PROGRAMS USED, AND TOTAL MINUTES OF ONLINE INSTRUCTION BY GRADE LEVELS AND SUBJECT AREAS

	NUMBER OF CORE ONLINE PROGRAMS USED (% OF OVERALL STUDENT SAMPLE)			
ELA	0	1	2	3
Grades K, 1 (N = 103)	3	89	8	0
Grades 2, 3 (N = 118)	6	36	58	0
Grades 4, 5 (N = 113)	3	4	20	73
Grades 6, 7, 8 (N = 79)	10	24	45	21
MATH				
Grades K, 1 (N = 103)	2	11	77	11
Grades 2, 3 (N = 118)	2	17	47	34
Grades 4, 5 (N = 113)	9	37	54	0
Grades 6, 7, 8 (N = 79)	13	22	41	24

TABLE 6: DESCRIPTIVE STATISTICS FOR INTENSITY OF USE OF ONLINE PROGRAMS, ELA (MINUTES)

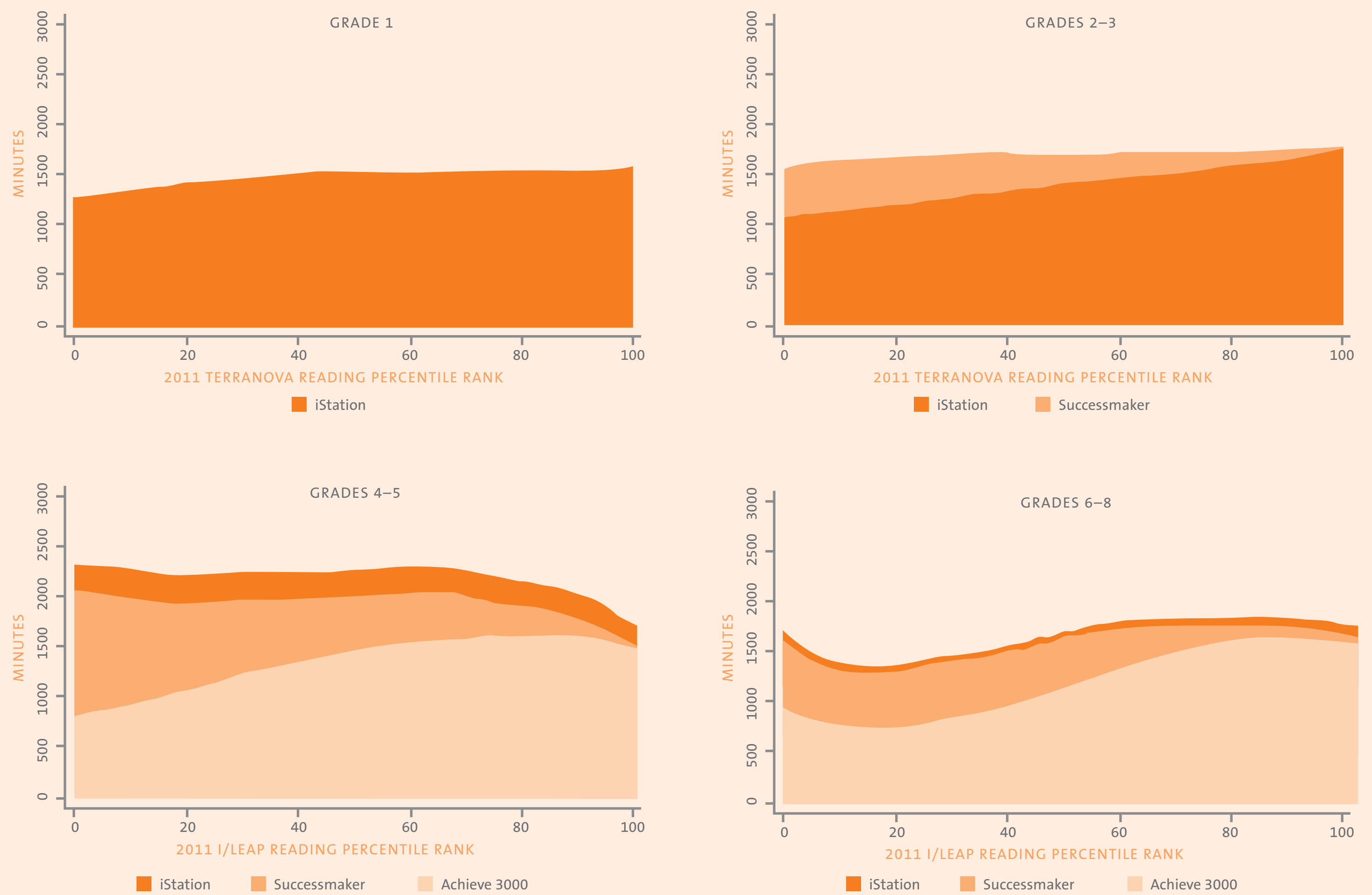
	% OF STUDENT SAMPLE USING PROGRAM	MEAN (MEDIAN)	STANDARD DEVIATION (MIN/MAX)
GRADES K, 1 (N = 103)		1,376 (1,411)	
iStation	97	1,376 (1,411)	274 (265/1,845)
GRADES 2, 3 (N = 118):		1,655 (1,709)	
iStation	94	1,369 (1,478)	447 (8/2,027)
SuccessMaker	58	467 (350)	461 (2/1,818)
GRADES 4, 5 (N = 113)		2491 (2,502)	
iStation	95	297 (300)	222 (9/807)
SuccessMaker	75	852 (968)	612 (2/2,009)
Achieve3000	93	1,617 (1,601)	920 (274/5,269)
GRADES 6, 7, 8 (N = 79)		1,988 (1,842)	
iStation	42	130 (152)	76 (3/267)
SuccessMaker	54	842 (590)	683 (4/2,256)
Achieve3000	77	1,618 (1,540)	895 (20/4,899)

TABLE 7: DESCRIPTIVE STATISTICS FOR INTENSITY OF USE OF ONLINE PROGRAMS, MATH (MINUTES)

	% OF STUDENT SAMPLE USING PROGRAM	MEAN (MEDIAN)	STANDARD DEVIATION (MIN/MAX)
GRADES K, 1 (N = 103)		1,403 (1,412)	
DreamBox	94	981 (1,073)	315 (349/1,623)
ST Math	91	494 (410)	279 (37/1,348)
GRADES 2, 3 (N = 118):		1,694 (1,589)	
DreamBox	68	604 (449)	417 (4/2,067)
ST Math	87	1,116 (1,081)	504 (71/2,477)
SuccessMaker	57	464 (271)	504 (1/1,738)
GRADES 4, 5 (N = 113)		2,027 (2,262)	
ST Math	88	1,516 (1,500)	977 (46/3,285)
SuccessMaker	57	892 (965)	644 (6/2,069)
GRADES 6, 7, 8 (N = 79)		2,270 (2,408)	
ST Math	76	1,182 (1,083)	820 (27/2,947)
SuccessMaker	61	1,119 (997)	872 (12/2,648)
Apangea	38	991 (916)	502 (344/2,690)

most frequently used for remediation purposes for both ELA and math, particularly with RTI Tier 2 and special education students in the learning support classrooms. In contrast, Apangea Math, an adaptive program with "live" online support that features multistep problem-solving was reserved for more advanced students in Grades 6–8.[26] Department leads, classroom teachers, and lab coaches made decisions as to which online programs were assigned to students, and assignments were fluid and changed across the school year. In addition to the total amount of time students spent on each online program by grade level, Tables 6 and 7 lists the percentage of students who used the core online programs in the different grade levels for ELA and math. The variation in percentages demonstrate how the school selectively used SuccessMaker across the different grade levels, as well as Apangea and iStation in Grades 6–8 and DreamBox in Grades 4 and 5 to personalize instruction to meet the needs of students with different levels of academic preparation.

FIGURE 1. DISTRIBUTION OF TIME OF USE OF CORE ONLINE INSTRUCTION PROGRAMS FOR ELA BY PROGRAM AND ACADEMIC NEED (PRIOR ACHIEVEMENT). (NOTE: PERCENTILES WERE GENERATED SEPARATELY FOR EACH GRADE LEVEL. THE PERCENTILES ARE NOT EQUIVALENT TO STATE NORMS.)



[26] *Apangea Math is now Think Through Math.*

FIGURE 2. DISTRIBUTION OF TIME OF USE OF CORE ONLINE INSTRUCTION PROGRAMS FOR MATH BY PROGRAM AND ACADEMIC NEED (PRIOR ACHIEVEMENT). (NOTE: PERCENTILES WERE GENERATED SEPARATELY FOR EACH GRADE LEVEL. THE PERCENTILES ARE NOT EQUIVALENT TO STATE NORMS.)

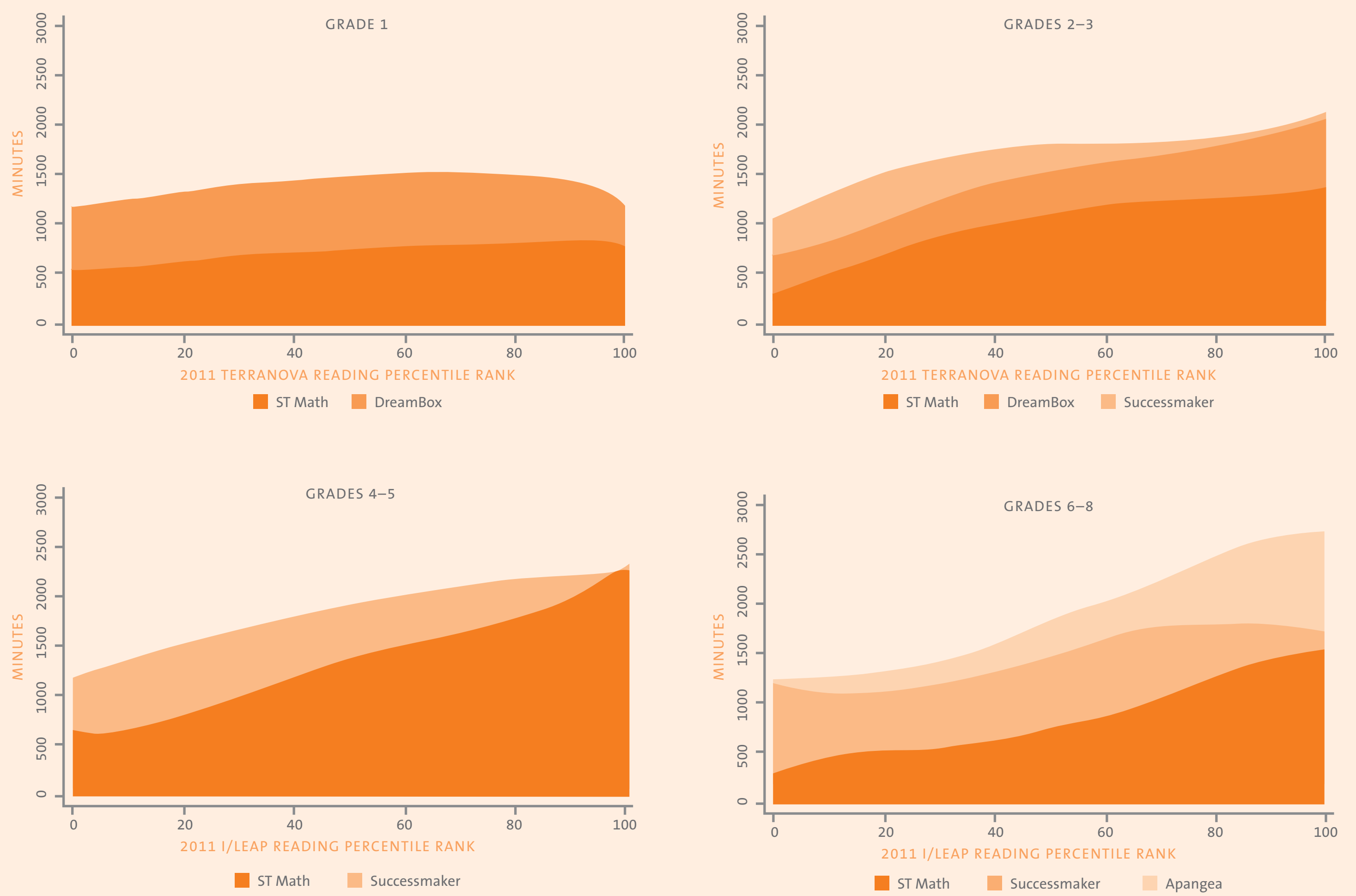


Table 5 shows the percentage of students who worked on one or more core online programs during the school year as well as the percentage of students who logged no minutes on these programs. For Grades K–1 and ELA, use of core programs was almost exclusively limited to a single program, iStation. For Grades 4–5, the majority of students
(77%), but not all, worked with three different programs (Achieve3000, SuccessMaker, and iStation). For Grades 2 and 3 and Grades 6 through 8 in ELA, there is greater variation across students within these grade levels in the number of programs accessed. In general, there is also greater within grade-level variation for math compared to ELA, with the exception of Grades K–1, where the majority of students (77%) used two different programs, DreamBox and ST Math.

Figures 1 and 2 further illustrate how the use of the different programs varied by grade level and academic need for ELA and math. The charts clearly show how the school used different combinations of programs in an attempt to personalize the learning experience for FirstLine students, particularly in Grades 2 through 8. For example, for online math instruction for Grades 6 through 8 (see Figure 2, lower right) students in the bottom 40% of prior achievement scores almost exclusively used ST Math and SuccessMaker for their online math instruction in the learning lab, while students in the top 40% of prior academic achievement increasingly were exposed to Apangea Math and ST Math, while fewer students were assigned to SuccessMaker. In

TABLE 8: AMOUNT OF TIME CORE ONLINE PROGRAMS USED BY RESPONSE TO INTERVENTION (RTI) TIER, ELA AND MATH (MINUTES)

	RTI TIER 1 MEDIAN MINUTES (N)	RTI TIER 2 MEDIAN MINUTES (N)	RTI TIER 3 MEDIAN MINUTES (N)
GRADES K, 1 (N = 103)			
Total ELA	1,435 (87)	N/A	1,264 (13)
Total Math	1,421 (87)	N/A	1,322 (13)
GRADES 2, 3 (N = 118)			
Total ELA	1,715 (89)	N/A	1,687 (22)
Total Math	1,671 (94)	N/A	1,325 (22)
GRADES 4, 5 (N = 113)			
Total ELA	2,281 (31)	2,531 (53)	2,612 (26)
Total Math	2,617 (27)	2,260 (52)	1,201 (24)
GRADES 6, 7, 8 (N = 79)			
Total ELA	1,702 (29)	2,556 (23)	1,517 (18)
Total Math	2,850 (30)	2,283 (23)	757 (15)

Grades 6–8, only 30 students ever logged time on Apangea Math to support their math instruction, and these students tended to have higher prior achievement scores than their peers.

Students with higher prior academic performance spent more total time in online math instruction. Students' academic needs appear to also determine the amount of time they experienced in online math instruction. The relationship between academic need, as measured by prior achievement scores, and total time was most pronounced for online math instruction. In Grades 2–3, students with lower prior achievement scores, spent more time on SuccessMaker than their higher-scoring peers, but overall, students spent as many as 500 fewer minutes in online instruction, or 25% less overall time (about 15 minutes per week). A similar trend exists in Grades 4–5, whereby students who scored higher on the 2011 state math achievement test logged as many as 900 more minutes (25 minutes per week) on ST Math and 400 more minutes overall. The largest difference in total time spent in online instructional environments between students with lower and higher prior achievement was in Grades 6–8 where students who scored in the 90th percentile on the 2011 state math exam logged 2,000 more minutes (55 minutes per week) than students who score below the 40th percentile.

Given that all students spent the same amount of time in the learning lab, including the learning-support classrooms, the discrepancy in the amount of online instruction between students with different levels of prior academic success in math is likely the result of differences in the amount of direct instruction these students received in the learning lab. FirstLine's blended learning model offers students with the greatest academic needs additional supports during their learning lab time including software that may be better suited for their needs than the core online programs as well as one-on-one and small-group direct instruction. Many of the lower-performing students likely spent a good percentage of their time—the DBL estimates about 50%—in the learning lab receiving some form of direct teacher-led instruction (one-on-one or small group) while, at the same time, their peers received instruction from the online programs. This is particularly evident in math in Grades 4 through 8. Table 8 shows the amount of time by grade level that students spent in online instruction by their RTI tier. The 24 Tier 3 students in Grades 4 and 5 spent half as much time working with the core online programs compared to their peers. The difference is even greater for the older students. In Grades 6 through 8, the 15 Tier 3 students only spent a quarter of the time as their peers working on the core online programs. The median Tier 3 students averaged about 20 minutes per week, compared to nearly 90 minutes per week for Tier 1 students.

Impact Analyses

IMPACT DESIGN OVERVIEW

A quasi-experimental design was implemented to explore the relationship between adoption of the FirstLine blended learning model and gains in student learning. The design involved comparing the learning outcomes for students enrolled in Arthur Ashe Charter School in the 2011–12 school year to comparison students attending one of three other elementary schools in the FirstLine charter school network. Since two of the three comparison schools experimented with blended learning beginning partway through the school year, the estimated impacts of the blended learning model adopted by Ashe based on this quasi-experimental design may underestimate the effects of the model (assuming that the blended learning models in Dibert and Langston Hughes also had positive effects on learning).

Michael & Susan Dell FOUNDATION

Because of the high level of missing data for the achievement outcomes in Grades 4–8, the impact analysis for FirstLine was limited to Grades 1–3. Students' scores on the spring 2012 Terra Nova assessment for students in Grades 1–3 were compared between Ashe and the other three FirstLine elementary schools, controlling for spring 2011 scores.

Descriptive statistics for Ashe students in school year 2011–12 and students enrolled in the comparison schools included in the analyses appear in Table A.1 in the Appendix for this profile. Ashe was relatively comparable to the comparison schools on all demographic variables, with slightly elevated levels of students with special academic needs as indicated by whether students had an Individualized Education Plan (IEP).[27] Of the students who were included in the analysis, approximately 20% of Ashe students included in the impact analysis had an IEP, compared to a range of 11% to 16% of students in the comparison schools.[28] Table A.2 shows the equivalence between the groups compared on prior achievement scores for the student samples included in the impact analyses. There were no statistically significant differences in prior math or reading Terra Nova scores between Ashe students and students in the comparison schools.

Tables 9–12 show the results for the impact analyses. Table 9 shows the estimated impacts for the overall student sample based on a comparison of students in Ashe to students in the other FirstLine schools. Impact estimates for selected subgroups (gender, prior achievement, and IEP status) are shown in Tables 11–14.

Limitations. Although the comparison schools may be similar in many ways to the schools using Khan Academy, there may have been important differences between Arthur Ashe and the comparison schools (e.g., differences in curriculum, academic culture, school leadership, and teaching staff) that may explain differences between the schools on spring test scores that are totally unrelated to the use of Khan Academy. Thus, using this design we could not isolate the effect of the introduction of Khan Academy from other key differences between schools that are likely to influence student achievement.

IMPACT FINDINGS

We found evidence of improved student learning in math for Grades 1–3 associated with attending Arthur Ashe in 2011–12 relative to other schools in the FirstLine network (Table 9). The effect for math in Grades 1–3 (effect size = +0.40) represents a difference of 22 scale points on the Terra Nova or a 4% gain relative to the average score in the comparison schools. It can also be interpreted as the median student in Grades 1–3 in the comparison schools (50th percentile) moving up 15 percentiles if that student had attended Arthur Ashe (from 50th to the 65th percentile). The size of the effect is considered moderate by educational research standards. For ELA, no substantive effects were found.

We found no evidence that the effects in Grades 1–3 for math or ELA varied by gender; however, the effects on Grades 1–3 math scores did

TABLE 9: OVERALL EFFECTS ON TEST SCORES (ASHE VS. COMPARISON SCHOOLS)

OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
TERRA NOVA (GRADE 1-3)									
Reading	137	602.1	29.1	465	599.5	44.6	2.7	0.06	0.425
Math	137	579.4	47.3	465	557.6	55.7	21.8	0.40	0.000

Note: Estimated impact is the coefficient from the model associated with attending Arthur Ashe. Effect size = estimated impact from model/pooled standard deviation (SD) between Ashe and comparison schools.

[27] *Individualized education plans (IEPs) are mandated by the Individuals with Disabilities Education Act. An IEP defines the individualized learning objectives of a child who has a disability, as defined by federal regulations, and any special supports and resources needed to help the child achieve those objectives.*

[28] *Overall, 40% of Ashe students had IEPs, with a majority of these students being older students in Grades 4–8. These grade levels were not included in the impact analyses due to the high level of missing achievement data.*

TABLE 10: EFFECTS ON TEST SCORES, BY GENDER (ARTHUR ASHE VS. COMPARISON SCHOOLS, SY 2011–12)

OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
TERRA NOVA (GRADE 1-3)									
Reading									
Female	64	601.4	30.1	209	602.0	37.7	-0.6	-0.02	0.897
Male	73	603.1	27.8	256	597.4	48.7	5.8	0.14	0.209
Math									
Female	64	574.4	42.8	209	556.8	53.1	17.6	0.33	0.008
Male	73	582.8	51.2	256	558.2	57.8	24.5	0.45	0.000

Note: Estimated impact is the coefficient from the model associated with attending Arthur Ashe. Effect size = estimated impact from model/pooled standard deviation (SD) between Ashe and comparison schools.

TABLE 11: EFFECTS ON TEST SCORES, BY STUDENT PRIOR ACHIEVEMENT (ARTHUR ASHE VS. COMPARISON SCHOOLS, SY 2011–12)

OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
TERRA NOVA (GRADE 1-3)									
Reading									
High Prior	59	603.6	27.3	228	606.7	32.9	-3.1	-0.07	0.544
Low Prior	78	600.8	24.4	237	592.6	44.4	8.2	0.20	0.067
Math*									
High Prior	56	567.9	44.0	243	562.5	49.1	5.4	0.10	0.429
Low Prior	81	586.6	47.8	222	552.2	50.9	34.4	0.64	0.000

Note: Estimated impact is the coefficient from the model associated with attending Arthur Ashe. Effect size = estimated impact from model/pooled standard deviation (SD) between Ashe and comparison schools.

**Indicates statistically significant differences in effect sizes between high prior achievement and low prior achievement groups, $p < .05$.*

vary significantly by levels of prior achievement and IEP status (see Tables 10-12). Specifically, the effect on math learning was greater for students with lower prior test scores (effect size = +0.64; statistically significant) compared to the effect for their higher scoring classmates (effect size = +0.10). The effect for students with lower prior test scores in Ashe represents a gain of 34 scale points on the Terra Nova relative to similar students in the comparison schools (a 6% gain) compared to a 5-point relative gain for their higher scoring peers (less than a 1% gain; see Table 11). The effect of attending Ashe on students with lower prior test scores is responsible for a majority of the overall effect on spring math test scores found for Grades 1–3. We also found differences in the effects on math scores in Grade 1–3 between students with and without an Individualized Education Program (IEP) (see Table 12). The effect on math learning for students without an IEP (effect size = +.45; statistically significant) was greater than the effect for students with an IEP (effect size = +.20). Students without an IEP outgained their peers in the comparison schools by 24 scale points (effect size = +.45) compared to an 11-point relative gain for Ashe students with an IEP (effect size = +.20). For ELA, the opposite was true. Students with an IEP outperformed students with an IEP relative to their peers in the comparison schools (effect sizes of +.37 and +0.04, respectively), but the difference in the effects between the two groups was not statistically significant.

Michael & Susan Dell FOUNDATION

TABLE 12: EFFECTS ON TEST SCORES, BY IEP STATUS (ARTHUR ASHE VS. COMPARISON SCHOOLS, SCHOOL YEAR [SY] 2011–12)									
OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	ESTIMATED IMPACT	EFFECT SIZE	P
TERRA NOVA (GRADE 1-3)									
Reading									
IEP	27	593.2	29.1	64	577.9	55.2	15.3	0.37	0.049
No IEP	110	604.5	26.1	401	602.9	39.6	1.6	0.04	0.654
Math*									
IEP	27	562.6	50.1	64	551.7	61.1	10.9	0.20	0.307
No IEP	110	582.9	43.4	401	558.5	53.6	24.3	0.45	0.000

Note: Estimated impact is the coefficient from the model associated with attending Arthur Ashe. Effect size = estimated impact from model/pooled standard deviation (SD) between Ashe and comparison schools.

**Indicates statistically significant differences in effect sizes between no IEP and IEP groups, p < .05.*

Appendix: FirstLine Schools – Arthur Ashe Charter School

TABLE A.1: DESCRIPTIVE STATISTICS: STUDENTS IN THE ANALYTIC SAMPLE FOR ASHE AND COMPARISON SCHOOLS (GRADES 1–3)				
VARIABLE	ASHE	DIBERT	GREEN	LHA
ETHNICITY				
American Indian/Alaska Native	0%	0%	0%	0%
Asian	1%	0%	0%	0%
Black/African American	98%	85%	96%	100%
Hispanic/Latino	0%	9%	2%	0%
White	1%	6%	2%	0%
NUMBER OF STUDENTS IN EACH GRADE				
1	54	59	54	66
2	58	35	51	72
3	62	46	66	76
GENDER				
Female	46%	36%	51%	48%
Male	54%	64%	49%	52%
ENGLISH LANGUAGE LEARNER (ELL) STATUS				
ELL	1%	8%	1%	0%
Not ELL	99%	92%	99%	100%
FREE/REDUCED-PRICE LUNCH (F/RP)				
F/RP Lunch	98%	96%	97%	98%
Paid Lunch	2%	4%	3%	2%
INDIVIDUAL EDUCATION PLAN (IEP) STATUS[a]				
IEP	20%	11%	16%	14%
No IEP	80%	89%	84%	86%

Note: Key to school names: Ashe = Arthur Ashe Charter School; Dibert = John Dibert Community School; Green = Samuel J. Green Charter School; LHA = Langston Hughes Academy
a. Includes special education and RTI Tier 3.

TABLE A.2: TEST OF BASELINE EQUIVALENCE: STUDENTS IN THE ANALYTIC SAMPLE BETWEEN ASHE AND COMPARISON SCHOOLS

OUTCOMES	TREATMENT			COMPARISON			CONTRAST		
	N	MODEL-ADJUSTED MEAN	SD	N	MODEL-ADJUSTED MEAN	SD	DIFFERENCE	ES[a]	P
TERRA NOVA (GRADE 1-3)									
2011 Reading	137	-0.1	0.9	465	0.0	1.0	0.0	-0.04	0.874
2011 Math	137	0.0	1.0	465	0.0	1.0	0.1	0.08	0.846

a. SD = standard deviation; ES = effect size

Teacher Survey Results – FirstLine (Arthur Ashe)

SURVEY ADMINISTRATION

Overall, 19 of 20 FirstLine teachers implementing blended learning at Arthur Ashe completed the Teacher Survey for a 95% response rate. Table S1 describes the characteristics of the respondents in terms of the subject area(s) they teach and years of teaching experience. Additionally, 2 of 2 lab monitors supporting the learning labs at Ashe completed a Lab Monitor Survey for a 100% response rate, and 47 of 72 teachers at nonblended FirstLine elementary schools completed the Comparison Teacher Survey for a 65% response rate.

TABLE S1. TEACHER SURVEY RESPONDENT CHARACTERISTICS

TEACHER CHARACTERISTICS	NUMBER	PERCENT
SUBJECT TAUGHT		
English language arts (ELA)	9	47
Math	7	37
Math and ELA	3	16
TEACHING EXPERIENCE		
New teachers (3 years or less)	9	47
Veteran teachers (4 or more years)	10	53

CLASSROOM INSTRUCTIONAL ACTIVITIES

Although the blended learning model implemented in the schools was primarily lab-based, there were distinct differences between Ashe and the comparison schools in how regular classroom instruction was organized (see Figures S1 and S2). Teachers in Ashe reported that

FIGURE S1: PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (ELA CLASSROOMS)

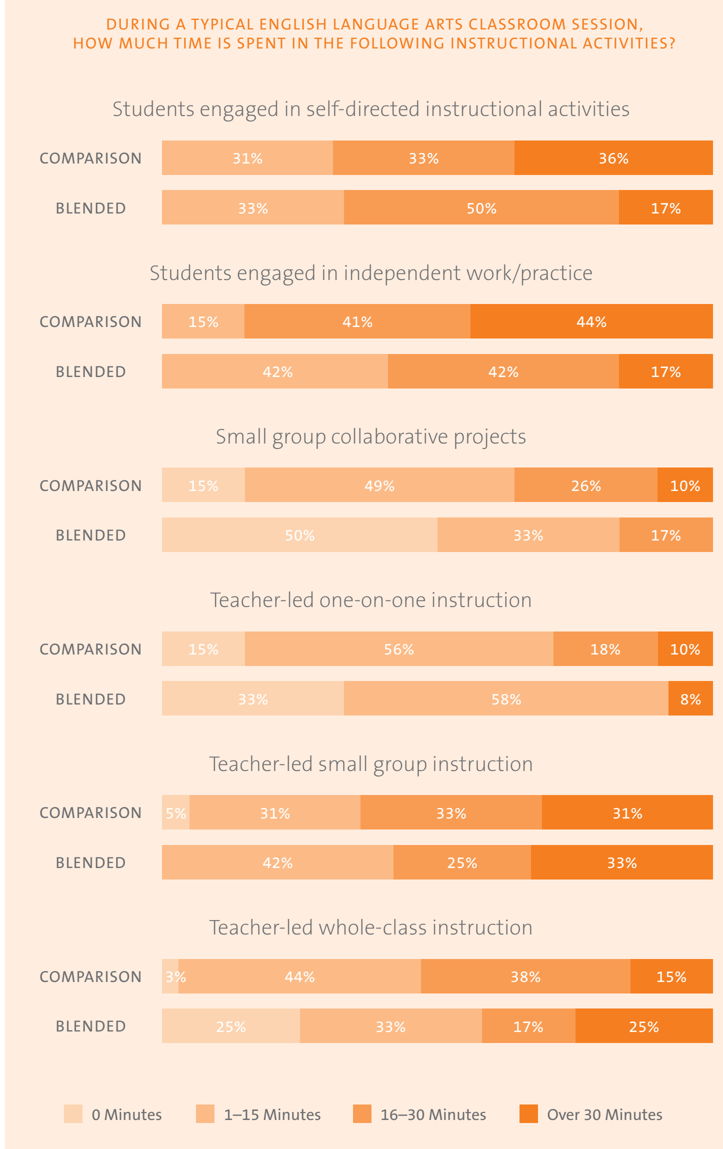
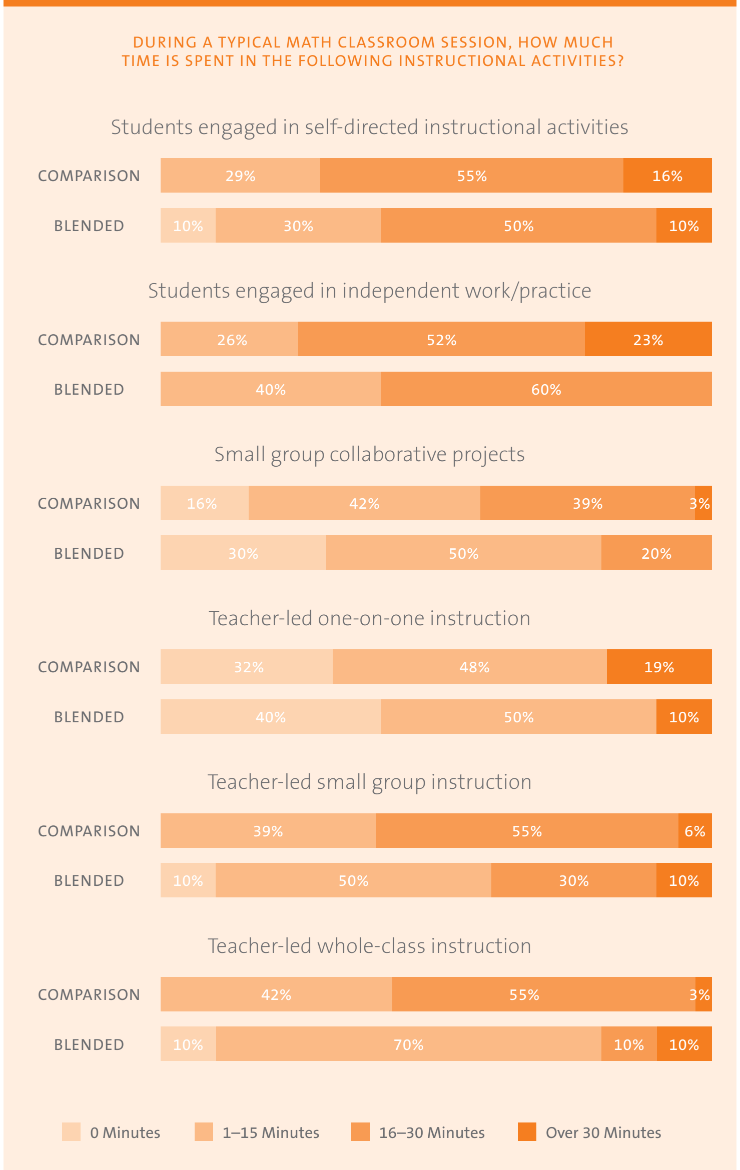


FIGURE S2: PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (MATH CLASSROOMS)



students spent less time in independent activities, as well as small-group collaborations and instruction. At the same time, students at Ashe spent less instructional time in whole-class teacher-directed lectures. The reason for these differences is not clear. Schools in the FirstLine networks had much discretion over curriculum, how instruction was organized, and school culture. Thus, the differences reported by teachers may represent local decisions made by administrators and staff about how to organize teaching and learning at their schools. However, it also may be the case that adoption of a blended learning lab model was shaping classroom instruction. For example, classroom teachers at Ashe used the learning lab time for small-group pull-outs and working with students one-on-one. In addition, a bulk of the time spent by students in the learning lab was spent working independently on the software programs. As a result, teachers may have spent less time in the classroom in these activities than they would without the adoption of the blended learning model.

ROLE OF TECHNOLOGY

Nearly three-quarters of teachers reported, that in addition to remediation/filling in knowledge gaps, technology and web-based instruction played a large role in personalizing instruction (e.g., allowing students to learn at their own pace, to meet the needs/interests of different types of learners), as well as capturing data on student academic performance (e.g., to capture data to inform instruction, for diagnostic or formative assessment). Fewer than one-third of teachers used technology or online instructional environments to introduce new concepts, and only 1 teacher in 10 used technology to support homework assignments (see Figure S3).

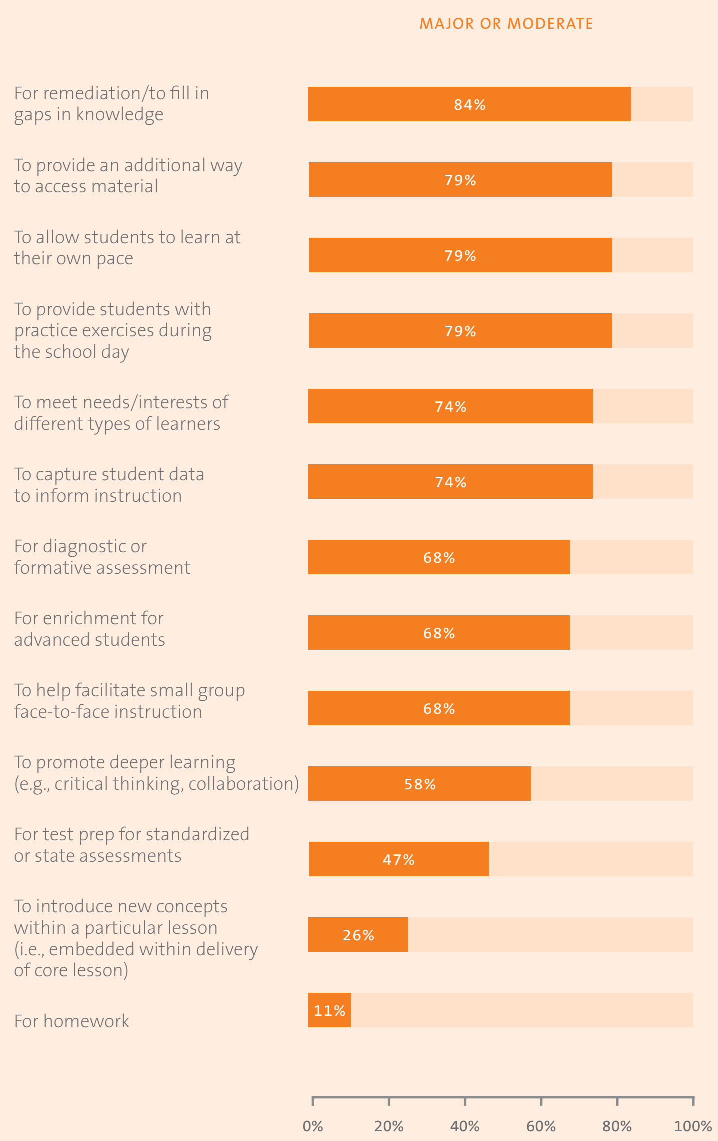
Technology and computer-mediated instruction played less of a role in several areas in math as compared to English language arts (ELA), including to capture data to inform instruction (43% vs. 89% in ELA), to help facilitate small-group instruction (43% vs. 78% in ELA), and to promote deeper learning (29% vs. 67% in ELA). Providing enrichment opportunities was the one area in which technology seemed to play more of a role in math than in ELA (86% vs. 56% in ELA).

THE LEARNING LAB

Lab monitors at Ashe were largely responsible for establishing and maintaining culture in the learning lab, rather than providing academic support, ensuring students were on task and engaged, and monitoring student progress within the software system(s). Many (69%) of FirstLine Ashe teachers said they supervised students in the learning lab at least once a week. However, more of the ELA teachers supervised students in the learning lab on a daily basis as compared to math teachers (44% vs. 14%, respectively). Most (79%) "pull students out" of the learning lab for individualized, remedial, or small-group instruction at least twice a week, and nearly half (47%) say

Michael & Susan Dell FOUNDATION

FIGURE S3. PERCENT OF TEACHERS REPORTING DIFFERENT ROLES FOR TECHNOLOGY AND WEB-BASED INSTRUCTION IN SUPPORTING TEACHER-PROVIDED INSTRUCTION



they do so every day or almost every day. In particular, math teachers reported running pull-outs on almost a daily basis, whereas many of the ELA teachers did this a few times a week, or even less frequently.

TABLE S2. ALIGNMENT OF CLASSROOM AND LEARNING LAB CURRICULUM

IN YOUR OPINION, HOW WELL ALIGNED IS THE MATERIAL STUDENTS ARE EXPOSED TO IN THE LEARNING LAB WITH THE MATERIAL YOU PRESENT IN THE CLASSROOM?	ELA (%)	MATH (%)	OVERALL (%)
Very well aligned	11	29	**16**
Moderately aligned	67	71	**74**
Weakly aligned	11	0	**5**
Not at all aligned	11	0	**5**
Don't know	0	0	**0**

More than 90% of teachers believed that the learning lab curriculum was moderately to very well aligned with the classroom curriculum (see Table S2). However, all (100%) of the math teachers indicated the material students were presented in the lab was moderately or very well aligned with the material they presented in the classroom, whereas only three-quarters (78%) of ELA teachers indicated the same.

Just over a third (37%) of teachers indicated that they met with lab monitors once a week or more to discuss students' performance in the learning lab; 42% did this three times a month or less, and 21% never meet with lab monitors. ELA teachers tended to meet with lab monitors to discuss student performance in the learning lab on a weekly basis, whereas math teachers tended to do so on a monthly basis. Frequently cited topics of discussion between teachers and lab monitors included students' progress within the software systems and students' behavior in the learning lab.

Most (84%) of teachers said it was very important or fairly important for them to know what their students were doing in the learning lab to be effective as their teacher, but close to half (42%) said they had some but not enough knowledge about what students struggled with in the learning lab, and 16% said they have no knowledge.

ELA teachers indicated it was more important for them to know what their students were working on in the learning lab to be effective in their role than the math teachers, and more of the ELA than math teachers indicated they had enough knowledge about what their students struggled with in the learning lab (44% vs. 29%, respectively); this may be why ELA teachers tended to stay in the learning lab to supervise students rather than run pull-outs. Overall, two-thirds (68%) of teachers said they would like to meet with lab monitors to discuss student performance in the learning lab more often than they did.

USE OF DATA

Nearly all (95%) Ashe teachers reviewed the student performance data recorded by the software system(s) used to support blended learning, and did so once a week or more (see Table S3). While the bulk of both math and ELA teachers reviewed data on a weekly basis, approximately a third (37%) of the ELA teachers reported doing so even more frequently.

TABLE S3. FREQUENCY OF DATA REVIEW BY FIRSTLINE TEACHERS

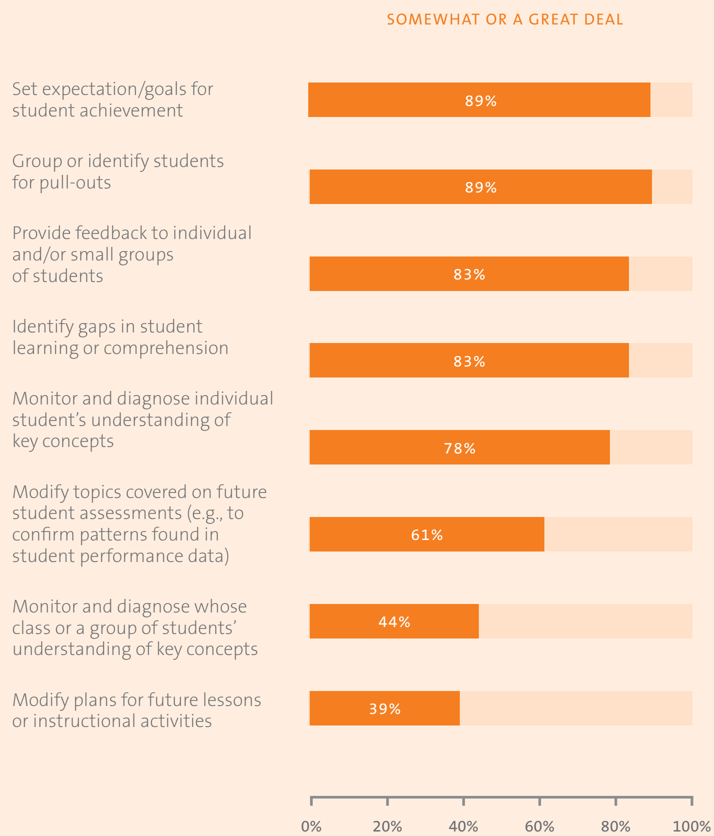
HOW OFTEN DO YOU LOOK AT THE DATA RECORDED BY THE SYSTEM(S) USED TO SUPPORT BLENDED LEARNING?	ELA (%)	MATH (%)	OVERALL (%)
Every day or almost every day	25	0	**11**
Two to three times a week	12	14	**22**
About once a week	62	86	**67**
Two to three times a month	0	0	**0**
About once a month or less	0	0	**0**

While nearly half (44%) found the system reports “very useful” for informing their instruction, fewer than 4 in 10 teachers say they used the data in the reports to modify future lessons or instructional activities. In fact, Ashe teachers’ review of the student performance data appears to influence their actions and decisions about individuals or small groups of students (e.g., set expectations/goals for student achievement, group or identify students for pull-outs, provide feedback to individual and/or small groups of students) more so than an entire class (e.g., modify topics covered on future student assessments, modify plans for future lessons or instructional activities) (see Figure S4).

ELA teachers indicated that the student performance data captured by the software program(s) was used to a greater extent than in math to modify future instruction plans and topics covered on future assessment, provide feedback to students, and monitor and diagnose students’ understanding of concepts. In math, the data were primarily used to identify gaps in student learning, group or identify students for pull-outs, and set expectations for student achievement.

FIGURE S4. PERCENTAGE OF TEACHERS USING DATA “SOMEWHAT” OR “A GREAT DEAL” FOR DIFFERENT INSTRUCTIONAL PURPOSES.

TO WHAT EXTENT HAS YOUR REVIEW OF STUDENT PERFORMANCE DATA FROM THE PROGRAM(S) LED YOU TO DO THE FOLLOWNG?



Satisfaction with Blended Learning and Impact on Student Learning

TEACHER SATISFACTION WITH BLENDED LEARNING

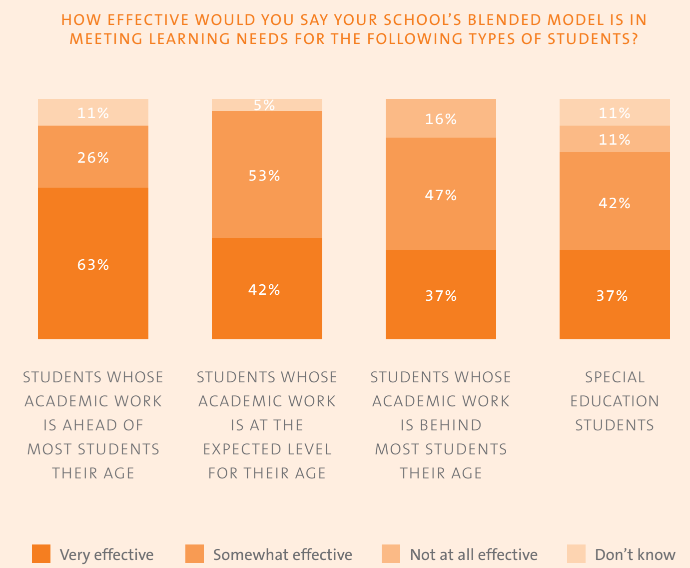
Almost all teachers reported they were satisfied with the blending learning model at Ashe and would recommend it to other teachers (see Table S4). However, although both math and ELA teachers generally appeared to be satisfied with the school’s blended model and its impacts on students, math teachers were even more positive in their response than ELA teachers.

Michael & Susan Dell FOUNDATION

TABLE S4: PERCENTAGE OF TEACHERS REPORTING THEY "AGREE" OR "STRONGLY AGREE" THAT BLENDED LEARNING BENEFITS THEIR STUDENTS IN DIFFERENT WAYS.

PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS.	PERCENT AGREE OR STRONGLY AGREE		
	ELA (%)	MATH (%)	OVERALL (%)
The learning lab model meets the learning needs of my students.	78	100	**90**
Students are highly engaged while using the learning lab model.	78	100	**90**
The learning lab model helps students take ownership for their own learning.	67	100	**85**
Students' learning and understanding of the materials has improved due to the use of the learning lab model.	78	85	**85**
I would recommend the use of the learning lab model to other teachers.	89	100	**95**

FIGURE S5. PERCENTAGE OF TEACHERS REPORTING DIFFERENTLEVELS OF THE EFFECTIVENESS OF THEIR BLENDED LEARNING MODELS FOR DIFFERENT PURPOSES.



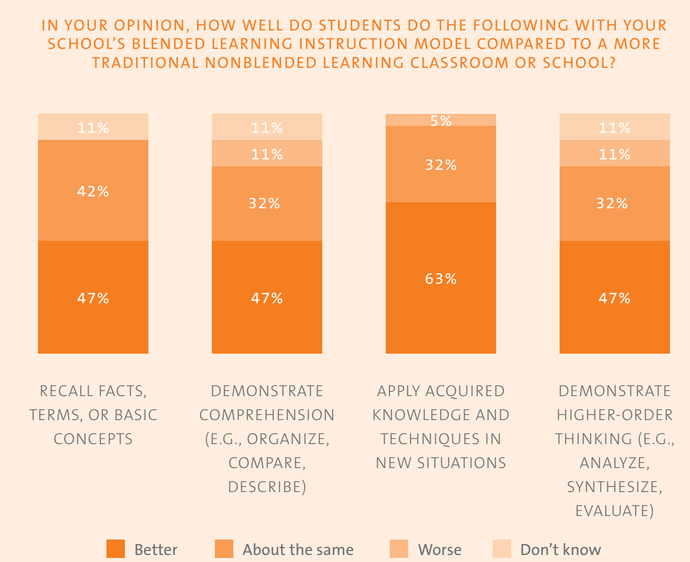
EFFECTIVENESS FOR STUDENT TYPES

Most Ashe teachers believed the learning lab model was at least somewhat or very effective for all students at various academic performance levels, particularly for students whose academic work was ahead of most students their age, though some teachers believed the model was not effective for students who were behind or in special education (see Figure S5).

At FirstLine Ashe, veteran teachers with 4 or more years of teaching experience tended to be more positive than new teachers with 3 or fewer years of experience when asked about students' skills under the school's blended model compared to a more traditional setting; close to two-thirds or more said students did better with blended learning, while half or fewer of the new teachers indicated the same.

Furthermore, most math teachers indicated the school's blended learning model was "very effective" for students whose academic work was ahead of or at the expected level for their age (86% and 71%, respectively), a belief shared by half or fewer of the ELA teachers. In contrast, ELA teachers reported blended learning was more effective for students whose academic work was behind most students their age than math teachers (56% said "very effective" vs. 14% of math teachers).

FIGURE S6. PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF THE EXTENT TO WHICH THEIR STUDENTS DEMONSTRATE HIGH-ORDER THINKING RELATIVE WITHIN THE BLENDED LEARNING MODEL RELATIVE TO MORE CONVENTIONAL INSTRUCTION MODELS



EFFECTIVENESS RELATIVE TO NONBLENDED ENVIRONMENT

Half or more of the teachers within Ashe reported that students demonstrated a variety of different skills better with blended learning than in a more traditional, non-blended environment, and few reported that their abilities in these areas were worse with the learning lab model (see Figure S6).

Veteran teachers with 4 or more years of teaching experience tended to be more positive than new teachers with 3 or fewer years of experience when asked about students' skills under the school's blended model compared to a more traditional setting; close to two-thirds or more said students did better with blended learning, while half or fewer of the new teachers indicated the same.

Math and ELA teachers, however, provided similar reports when asked about students' ability to recall facts, terms, or basic concepts with blended learning compared to a more traditional environment. However, math teachers were more positive than ELA teachers about students' ability to apply knowledge in new situations (71% report students did this better with blended learning compared to 44% of ELA teachers), and more divided in their opinions related to demonstrating comprehension and higher-order thinking.

TRAINING AND SUPPORT

Almost all (95%) of teachers participated in a training or orientation session (either in person or online) directly related to their school's use of blended learning or on the specific software program(s) supporting it. Close to two-thirds (62%) were satisfied or very satisfied with the training they received, while a third (39%) were dissatisfied or very dissatisfied. ELA teachers were more satisfied than math teachers with the training they were provided. Math teachers in particular wanted more information about accessing and interpreting student performance data captured by the software program(s).

Overall, about two-thirds of teachers spent 5 hours or less of their own time getting acquainted with the software program(s) or planning for how to best integrate blended learning with their instruction, while a third spent 6 or more hours (see Table S5). Additionally, although roughly half of both ELA and math teachers reported spending 1–5 hours of their own time preparing to implement blended learning, a third of the ELA teachers reported spending less than an hour of time, whereas close to half the math teachers reported spending more time.

TABLE S5. TIME SPENT OUTSIDE OF PROFESSIONAL DEVELOPMENT PREPARING FOR IMPLEMENTATION

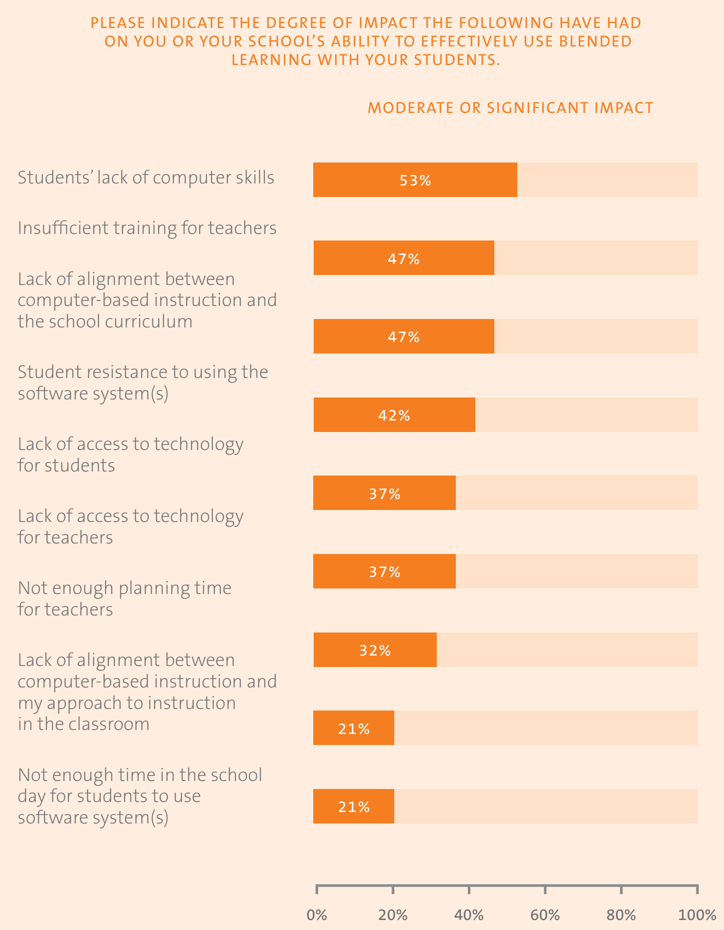
HOW MUCH OF YOU R OWN TIME (I.E., OUTSI DE OF ANY PROFESSIONAL DEVELOPMENT) DID YOU SPEND GETTING ACQUAINTED WITH THE SOFTWARE PROGRAM(S) SUPPORTING BLENDED LEARNING OR PLANNING FOR HOW TO BEST INTEGRATE BLENDED LEARNING WITH YOUR INSTRUCTION?	ELA (%)	MATH (%)	OVERALL (%)
Less than 1 hour	33	0	**16**
1–5 hours	44	43	**47**
6–10 hours	22	29	**26**
11–20 hours	0	14	**5**
21–30 hours	0	0	**0**
Over 30 hours	0	14	**5**

FACTORS INFLUENCING USE

Almost one-half of all teachers reported students' lack of computer skills, insufficient training for teachers, and lack of alignment between computer-based instruction and the school curriculum as the most prevalent nontechnology factors that had a moderate or significant degree of impact on the school and teachers' ability to effectively use the learning lab model at Ashe. In addition, almost 4 in 10 teachers reported that students' resistance to using the software programs was also a moderate to significant factor hindering the effectiveness of the blended learning model (see Figure S7). Consistent with their dissatisfaction with the training provided, a higher percent of math teachers indicated that insufficient training had a moderate or significant impact on their ability to implement blended learning effectively (71% vs. 33% of ELA teachers). Across the subject areas, teachers were otherwise relatively consistent in terms of other reported barriers.

All (100%) teachers reported that they or their students experienced technical problems in using the learning lab model. The most commonly cited technical problems included Internet connection unreliable/down (95%), insufficient bandwidth/Internet too slow (79%), student computers not working properly (74%), and software glitches or bugs (74%). Half (53%) indicated the technical challenges were a "moderate" or "significant" barrier to their effective use of the model.

FIGURE S7. PERCENTAGE OF TEACHERS REPORTING THAT DIFFERENT FACTORS INFLUENCED EFFECTIVE USE OF THE LEARNING LAB MODEL TO A "MODERATE" OR "SIGNIFICANT" DEGREE



Michael & Susan Dell FOUNDATION

Profile: KIPP Empower Academy (2011–12 School Year)

Introduction

KIPP LA Schools (KIPP LA), a nonprofit charter management organization (CMO), operates high-performing, free, open-enrollment public charter schools serving over 3,000 elementary and middle school students in South and East Los Angeles (http://www.kippla.org/). KIPP LA is part of the national KIPP (Knowledge Is Power Program) network of 141 schools in 20 states and Washington, D.C., whose mission is to prepare students in underserved communities for success in college and in life. KIPP LA currently operates nine schools in South and East Los Angeles, including five middle schools—KIPP Academy of Opportunity, KIPP LA Prep, KIPP Philosophers Academy, KIPP Scholar Academy, and KIPP Sol Academy —and four elementary schools, including KIPP Raíces Academy, KIPP Comienza Community Prep, KIPP Empower Academy, and KIPP Iluminar Academy. Nearly all (99%) of KIPP LA's students are African American or Latino, and the majority (89%) are low-income.

KIPP Empower Academy (KEA) was the first school in the KIPP LA region to implement a comprehensive blended learning model. KEA first opened in the 2010–11 school year, and during the study year (2011–12) KEA enrolled students in kindergarten and first grade. The school will eventually serve grades K–4. The school's instructional model was originally designed around blended learning. The founding principal chose to adopt a blended learning approach to enable teachers to deliver instruction to small groups of students as a means to differentiate instruction within mixed-ability classrooms.

BLENDED LEARNING MODEL DESCRIPTION

KEA adopted a rotational blended learning model within classrooms to support differentiated small-group instruction. Students rotated between teacher-led instructional activities, independent work, and online instruction within the classroom. When not working with teachers, students worked independently with adaptive online software programs, progressing at their own pace with occasional support from adults. For most of the class period, teachers led small-group activities. At the same time, a technology specialist floated between classrooms, monitoring students while they worked on the computers and addressed any technical problems that may arise. During the 2011–12 school year, the online programs KEA primarily used for kindergarten and first grade were iStation for English language arts (ELA) and DreamBox for math. Although the online programs used were selected based on their alignment with California Common Core State Standards (CCSS), the online content had not been mapped or sequenced to the school's offline curriculum.[29] The rotation model used varied by subject area and grade level.

English language arts instruction. During the 2011–12 school year, ELA instruction for kindergarten students was organized into three stations, two in which teachers led instruction—phonics/fluency and reading comprehension—and one targeting skill building where students received instruction through an adaptive software program. KEA used iStation to support ELA instruction in kindergarten and first grade. Groupings varied in size (6–14 students) and were "leveled" by academic ability based on performance on the school's benchmark assessments administered five times a year. In first grade, there were four stations—online instruction, teacher-led guided reading, independent reading, and independent work in workbooks. Two teachers were always present in each classroom during ELA. In kindergarten, an experienced lead teacher led the phonics/fluency lesson and a less experienced ("intervention") teacher, led the guided reading station.[30] Each station was 20–30 minutes long with students rotating through all stations in a 90-minute block. Every week, students received approximately 90 minutes (first grade) to 2 hours and 30 minutes (kindergarten) of online instruction in ELA.

Math instruction. In the 2011–12 school year, the timing and structure of the online math instruction varied by grade level. In kindergarten, students received teacher-led instruction within their 45-minute math block and online math instruction within a 50-minute science block. Kindergarten students spent half their science block in teacher-led small group instruction and the other half working on an online math program. Kindergarteners could have received up to 2 hours of online math instruction each week. One to two teachers and an instructional assistant were present in kindergarten classrooms. DreamBox, an adaptive program that is designed to deliver content at the appropriate level for each child based on a diagnostic assessment, was the primary software program KEA used at the time.

In first-grade math classrooms, a single teacher was responsible for all instruction. The use of the computer-assisted learning station varied by classroom and students' achievement levels and was more fluid and less systematic than in kindergarten. There were approximately 100 students per grade level and at any one time during the school day, 50 students were receiving math instruction within a daily 60-minute math block. The 50 students were distributed across four classrooms led by two lead teachers and two intervention teachers. Each math group was leveled by ability: one low-level, two mid-level, and one high-level. Typically, for the low-level math group, students spent most or all of the class period receiving small group teacher-led instruction. However, the online instruction and computer-based instruction used could vary significantly from class to class and day to day at the discretion of the teachers. In one classroom that was observed in the spring, two different teachers taught two groups of students of different abilities—one middle-level and one low-level—in the same classroom. The teachers in the room had worked out an arrangement that allocated computer use by ability group. The computers at the back of the room with access to DreamBox were reserved for students in the low-level math group, after completing their in-class assignments. For the other students in the classrooms—those performing at grade level—iPads were available, which the teacher typically allowed students to use in the latter half of the class period if there was a particular application with appropriate content connected to the daily lesson. Students in first grade could spend up to 2½ hours per week in online math instruction.

[29] *A detailed description of the KIPP Empower Academy's instructional and operational model is available in a case study developed as part of this research study and published on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).*

[30] *For the 2012–13 school year, the school had only one teacher in the room in first grade.*

Michael & Susan Dell FOUNDATION

TABLE 1. KIPP EMPOWER ACADEMY DESCRIPTIVE STATISTICS (2011–12 SCHOOL YEAR)							
SCHOOL NAME	GRADE LEVELS	STUDENTS (#)	% LATINO HISPANIC	% AFRICAN AMERICAN	% ELIGIBLE FOR FEDERAL LUNCH PROGRAM	% SPECIAL EDUCATION	2011–12 API SIMILAR-SCHOOL RANKING[A]
KIPP Empower Academy	K-1	230	10%	89%	91%	8%	8%

[A] *As an elementary school serving only Grades K–1 in 2011–12, the school did not take the California state achievement tests, which start in Grade 2.*

Sample

During the 2011–12 school year KEA was one of three elementary schools in the KIPP LA region. It is located in South Los Angeles. KEA is one of the highest-performing schools in the KIPP LA region as well as in the KIPP national network.[31] Table 1 summarizes the school's student characteristics for the 2011–12 school year.

Data Collection

SITE VISITS

One member of the research team conducted a site visit to KIPP LA in October of 2011 and again in April of 2012. The October visit included: interviews with the KIPP LA leadership team and staff; interviews with the principals at KEA and another KIPP LA elementary school that had not adopted a blended model; and classroom observations in both KEA and the other school. The April site visit consisted of four classroom observations, interviews with school principals and two KIPP LA leadership team members, and five teacher interviews. Table 2 summarizes the activities conducted during the two site visits.

TEACHER SURVEY

Overall, seven of eight KIPP Empower Academy teachers implementing blended learning completed the teacher survey for an 88% response rate. Five of the teachers had 4 or more years of teaching experience while two had 3 years or less.

SYSTEM USE DATA

This report summarizes the results of an analysis of the adaptive DreamBox software-program use data for KEA students during the 2011–12 school year. We accessed system use data for both iStation (ELA) and DreamBox (math) for all students in kindergarten and first grade. However, since the data from iStation did not include intensity of use and performance measures that were comparable to the other online programs analyzed as part of this research, we excluded iStation data from the system use data analysis.[32] For the analysis of system use data for DreamBox, we examined both total time-on-task (intensity of use) and the number of lessons completed (progress).

TABLE 2. DATA COLLECTION SUMMARY, FALL 2011 AND SPRING 2012 SITE VISITS		
	FALL 2011 SITE VISIT	SPRING 2012 SITE VISIT
OBSERVATIONS		
Schools	2	1
Classrooms	6	4
INTERVIEWS		
CMO Leaders	2	2
School Leaders	1	1
Teachers	0	5

[31] *See KIPP Empower Academy's 2011 Report Card (http://www.kipp.org/00/docs/KIPP_ReportCard_2011/KIPP_2011ReportCard_Los_Angeles.pdf)*

[32] *Compass Learning was also used early in the school year for ELA and math instruction but was eventually dropped by the school sometime in October.*

Data on students' use of DreamBox was provided by the vendor. To develop a measure of intensity of use, the number of minutes that a student logged for each session—from log in to log out—were summed across the entire year. The time students spent working in DreamBox lessons represented the majority of the time students spent working on academic-related activities within DreamBox.[33] A measure of progress or lessons "passed" was developed using data provided by the vendor. Each lesson that a student started was labeled as "fail," "abandon," or "pass." All lessons marked "pass" were summed for each student over the entire year.

STUDENT LEARNING OUTCOMES

The learning outcome analyzed was individual student scores on the Northwest Evaluation Association's (NWEA) Measurement of Academic Progress (MAP) assessments in ELA and math. The NWEA MAP is a computer-based adaptive assessment that is aligned with the California's CCSS(http://www.nwea.org). KIPP LA's schools (and most schools in the KIPP national network) administer the NWEA MAP assessment each spring and in the fall for new students entering the schools. NWEA MAP scores were provided by KIPP LA, and NWEA provided MAP scores for a "virtual comparison group" (VCG) created for the research team using NWEA MAP scores for similar students in similar schools who took the MAP during the same test window (for additional details on the VCG analysis see the description in the main body of the report under Analytical Models section, page 17).

Implementation Findings

Findings from the site visits and teacher survey were used to help understand the facilitating factors and barriers to implementing the rotational blended learning model at KEA during the 2011–12 school year. Additionally, an analysis of the DreamBox system use data was used to help better understand the intensity with which students used DreamBox to support math instruction and the potential connection between intensity of use and improved student learning.

FINDINGS FROM THE SITE VISITS AND TEACHER SURVEY

Data from the fall 2011 and spring 2012 site visits and the teacher survey were used to report on different aspects of implementation of KEA's rotation model. These findings are discussed below by theme, including infrastructure and technology issues, the support and training teachers received, characteristics and quality of system software, and teachers' use of the system data to support their instruction. In addition, we report on teachers' overall satisfaction with the blended learning model and the benefits to students and teaching they perceive in the model. A summary of the survey results is included in the Appendix for this profile.

INFRASTRUCTURE AND TECHNOLOGY

Introducing iPads into the classroom provided teachers with more flexibility in the choice of instructional activities. To support the school's blended learning model within classrooms, each classroom had approximately 15 computers. In January 2012, the school received a grant to purchase iPads. The iPads were used primarily for independent practice using web-based math applications, many of which allowed teachers to assign content (e.g., BrainPOP, Motion Math: Hungry Fish). During the second semester, some teachers allowed their students the choice to use DreamBox or applications on the iPad during their computer-assisted learning rotation. At least one first-grade math teacher, who worked with students who were below grade level, used time on the iPads at the end of class as a reward for students who had made sufficient progress on their in-class assignments. One drawback of the iPads was that teachers did not have access to any cumulative data on students' progress within a specific application.

Few technical issues impacted use of online instruction.
Although KIPP LA leadership reported significant bandwidth issues that had to be resolved prior to opening the school in fall 2010, KEA teachers completing the survey reported very few problems related to technology that significantly impacted the implementation of the blended learning model in school year 2011–12. Although the majority (86%) of responding teachers indicated they or their students had experienced technical problems, none of the teachers indicated

[33] *Students also spent time in the Carnival section of the DreamBox program. In this section of the program students engaged in video game-like activities of two types: one that is content free but designed to develop problem-solving skills and the other designed for "pure fun" according to the publisher (http://www.dreambox.com/online-math-games-for-kids). Students spent between 2% and 10% of their total time on DreamBox in these activities.*

they were any more than a "slight" barrier to the effective use of blended learning in the school. The most commonly cited technical problems included software glitches or bugs (83%), unreliable Internet connections (67%), insufficient bandwidth (67%), and student computers not working properly (67%).

Single sign-on portal was used to support efficient transitions to the online instruction station. To access the online adaptive software programs, very young children needed to be able to quickly log in and out of the various online programs with no assistance from adults. During the 2010–11 school year KEA worked with Education Elements to develop the Student Launch Pad that allowed students to log in and out efficiently and access all programs through a single portal. The portal included an authentication system that allowed students to securely enter each of the programs by clicking on the appropriate series of images (picture of the teacher, the student, and the student's picture password) rather than entering complex alphanumeric usernames and passwords. The desktop displayed a set of icons representing the various software programs, allowing students to enter a particular program by clicking the appropriate icon.

Blended learning project manager played an important role in the school's implementation of blended learning. KIPP LA created a project manager role to coordinate the implementation of the blended learning model at KEA. The project manager was responsible for negotiating with and monitoring external vendors and consultants; working with the director of technology to establish technology requirements and procure hardware and software for the school; and tracking progress relative to the project timelines. This role freed up the principal from these tasks and allowed for more streamlined procurement and improved overall project management. KIPP LA leadership and the principal cited strong project management as one of the most critical factors supporting their implementation of blended learning.

TRAINING AND SUPPORT

Teachers were satisfied with the training they received on the online programs. All but one of the teachers at KEA participated in a training or orientation session provided by the online program vendors, and nearly all were satisfied with the training they received. Teachers reported spending very little of their own time outside of any professional development (5 hours or less) getting acquainted with the software programs. According to the principal, because the selected online software programs were designed to be used by students independent of the off-line classroom curriculum, the demands on teachers were low, as was the need for more intensive professional development. This is expected to change as the school adds higher grade levels and considers investing in software programs like ST Math that will allow teachers, in a limited way, to assign content in the online station so that the content students work on is more closely aligned with the current lesson in the curriculum.[34]

Summer school was used to prepare students for in-class rotations through instructional activities. A critical factor for the success of the school's rotation-station model was the ability of students to seamlessly transition from one station to the next. Smooth transitions allowed the teachers and students to maximize the amount of instruction possible within each of the stations. During the second week of KEA's regular 2-week "summer school," students repeatedly practiced the ritual of transitioning from station to station following a typical rotation schedule within ELA and math. Students practiced timed transitions, according to the principal, with a goal of moving from one station to the next in under 60 seconds to maximize instructional time. During the first 2 weeks of the school year, teachers also spent time modeling the different behaviors expected when students were in the online instruction station, including how to sit, where to place their hands, how to care for the headsets, as well as the procedures for logging in and out. Students were told they were not allowed to talk to the students on either side of them while working on the online programs. During the research teams' two site visits, students were observed on several occasions rotating on and off an online instruction station efficiently and quietly and were settled and ready to work within 60 seconds or less after leaving the previous station.

School-based IT support staff member was available to assist students with technical problems, freeing up teachers to focus on their own instruction and students in their stations. The technology specialist supported students when they encountered technology-related problems while they were in the online instruction station. The specialist roamed between classrooms throughout the day and

[34] *During the 2012–13 school year the school used ST Math in second grade, but it runs in adaptive mode and teachers were not assigning content.*

ensured that all hardware and software programs were working properly and helped students if they were having difficulty logging in or experiencing problems with the hardware or software. The specialist also coached students on which keystrokes to use to resolve common glitches in the software on their own (e.g., screen freezes) in the event the specialist was not in the room.

SOFTWARE SELECTION, AVAILABILITY, AND DESIGN

The school continually evaluated their online program offerings. The school evaluated more than 20 online programs using a set of criteria it developed, including whether the programs were aligned to the scope and sequence of curriculum; if they were sufficiently broad and deep in terms of content available for each grade level; and if they were age-appropriate, interactive, user-friendly, cloud-based, affordable, supported data analysis, and compatible with the school's learning management system. In the beginning of the 2011–12 school year, the school selected Compass Learning for math and ELA. According to the school's principal and teachers interviewed, by the research team's spring visit the school had stopped using Compass Learning due to issues with low student engagement while students were working with the program, which led to disruptive student behaviors. Because of this, the school pivoted midyear and adopted iStation for ELA instruction and DreamBox for math as their primary online instruction programs. Subsequently, the school planned to replace iStation with iReady for school year 2012–13. According to the principal, iReady provides students with immediate feedback on their progress (a feature that iStation lacks), which he believed would contribute to better student engagement and self-efficacy.[35]

There is a limited connection between on-line and off-line classroom instruction. The school's primary purpose for adopting blended learning was to facilitate differentiated small group teacher-led instruction. As a result, during the 2011–12 school year, there was no attempt to align the content that students worked on during the online instruction station with the content teachers were covering in the small-group instruction station. Only one of the seven teachers responding to the survey reported that the lack of alignment in curriculum and instruction between the online programs and teacher-provided instruction had a moderate or significant impact on their ability to effectively use blended learning with their students. The school specifically selected programs that were adaptive and allowed students to proceed at their own pace as a way to provide additional differentiated instruction within the classroom to complement the leveled small-group direct instruction. Because these adaptive software programs allowed students to take different learning paths that challenge each individual, students were not working on the same topic at the same time, or the same topic that the teacher covered during small-group instruction. The school did evaluate different online programs that allowed teachers to assign students specific content to work on, including Compass Learning, but at the time, they were unable to identify any program that had sufficient content coverage for Grades K–1 and that was also affordable.

While the specific content students worked on during the online instruction station for a given day or week may not have be aligned with what the teachers were presenting in small groups, the content of the online programs used was aligned with state content standards. As a result, the principal was confident that students were getting exposure to online content that was relevant to the state content standards.

USE OF DATA TO INFORM INSTRUCTION

There was limited use of the system use data by teachers to inform their instruction. Although most (86%) of the teachers reviewed the student performance data recorded by the online programs, the majority of those that did review the data reviewed it about once a month or less. Most teachers reported on the surveys that the reports were "somewhat" or "slightly" useful for informing their instruction, but often relied more on information outside the online systems such as results from their own quizzes, end of unit assessments, and standardized assessments administered by the school at different points during the school year (e.g., NWEA MAP assessments and the STEP literacy benchmark assessments). One first-grade teacher reported in the spring that she had not reviewed any data from any of the systems, while another teacher reported she had never seen a report of students' progress for DreamBox. The principal also reported that the teachers did not necessarily trust the online assessments and algorithms the software programs used to evaluate student proficiency in a particular subject area. According to the principal, teachers had more confidence in their own observations and formative assessments.

[35] *According to KIPP LA leadership, the school did switch to iReady in August 2012 but switched back to iStation in January 2013.*

While KEA teachers used other sources of student performance data regularly to inform their instruction, the use of data from the online programs was seriously curtailed due to a lack of standardization across online programs in the type of data that was available to the principal and teachers. Because of this lack of standardization, it was not possible for the school principal and teachers to easily make judgments about student performance and progress based on the data provided by the multiple systems. During the 2011–12 school year, the school worked with Education Elements to develop an integrated solution to this problem, but by the end of the 2011–12 school year a usable teacher "dashboard" was not available.

SATISFACTION WITH BLENDED LEARNING AND PERCEIVED BENEFITS FOR STUDENTS

Teachers were satisfied with the use of blended instruction. KEA's teachers were generally positive in their survey responses about their school's blended learning model, with all teachers agreeing that they would recommend the use of the in-class rotation model to other teachers, that students were highly engaged while working at the online stations, and that the blended learning model met the learning needs of their students.

School staff believed the blended model was improving student learning and computer literacy skills. Teachers reported that they believe the rotation blended learning model was at least somewhat effective in improving learning for all groups of students, regardless of their academic preparation. And according to the principal, students' time on the online programs was also improving their computer literacy skills. He commented that many of the students enrolled in the school were living in homes with few computers relative to more affluent communities. As a result of students' exposure to computers in the school and the development of computer literacy skills, the principal believed students should be better prepared for the Common Core assessments since these are to be administered online.

School staff believed that the use of technology was effectively supporting their instruction. All KEA teachers responding to the survey indicated that technology-based instruction played a "major" or "moderate" role in supporting the instruction in different ways, including allowing students to learn at their own pace; filling in gaps in student knowledge; offering enrichment for advanced students; providing students with additional practice; and capturing data to help inform instruction. Six of the seven teachers also noted that use of technology-supported instruction played an important role in helping teachers facilitate small group face-to-face instruction (this was also emphasized in our interviews with teachers). In contrast, only two of the seven teachers believe that the technology-supported instruction promoted the acquisition of conceptual knowledge ("deeper learning of concepts").

According to the principal, in the spring students' work on the online programs was also helping the teachers prepare students for the state tests in two ways. First, it did so by allowing teachers to focus their ELA and math instruction on the skills students needed to work on while allowing teachers to pass over topics they knew were covered particularly well in the online programs. In addition, in math, the principal reported that the school relied on DreamBox to plug some gaps in content coverage in the core offline curriculum—Singapore Math.

FINDINGS FROM INSTRUCTIONAL SYSTEM LOG DATA

As described above, the analysis of system use data was limited to the student data available from DreamBox. To capture the intensity with which students used DreamBox, we computed the median number of minutes that students spent working on DreamBox lessons—the core instructional component of the program—over the course of the school year and the average number of minutes students worked on lessons during a median session. This was computed separately for kindergarten and first grade (see Table 3).[36]

TABLE 3. INTENSITY OF USE AND PERFORMANCE INDICATORS BY GRADE

	N	MEDIAN	MINIMUM	MAXIMUM
MINUTES (INTENSITY OF USE)				
Kindergarten	96	1,905	1,212	2,620
Grade 1	106	2,178	891	4,173
LESSONS COMPLETED (PROGRESS)				
Kindergarten	96	115	38	216
Grade 1	106	133	38	323

[36] *Students also used the program iStation for ELA. However, since the data provided from iStation did not include time-of-use information or performance data that was comparable to other programs used by schools in the study, the use of iStation was not included in this analysis.*

Time-on-task and number of days used. The median number of minutes spent on DreamBox for kindergartners was 1,873 (about 31 hours or 52 minutes per week for a 36-week school year) and ranged from of 1,155 to 2,576 minutes (19 to 43 hours). For students in first grade at KEA the median number of minutes logged by students was 1,946 (about 32 hours or 54 minutes per week) and ranged from 754 minutes to 4,003 minutes (13 hours to 67 hours). The maximum amount of time per week allotted by the model for online math instruction was 2 hours in kindergarten and 2½ hours in first grade. Discrepancies between the maximum amount of time available for online instruction during the school year and the amount of time spent on DreamBox computed for the median student may be explained by a host of factors, including in first grade, teachers had discretion over the amount of time allotted to the online instruction station, particularly for struggling math learners; students spent a portion of their time in DreamBox on nonlesson activities; some first-grade students were allowed to use iPad apps in addition to DreamBox; and the school's use of Compass Learning early in the school year to supplement DreamBox.

Student progress. To understand variation in students' progress within DreamBox, we analyzed the number of "unique lessons" that a student successfully completed. A lesson in DreamBox is comprised of multiple activities organized around a topic area, such as counting. DreamBox offers over 720 lessons across six grades, K–5. The length of time required to complete a lesson varies based on the difficulty of the content. Activities within a lesson are often variations on a common problem-type and a student must complete a set number of activities successfully in order to have "completed" the lesson. Thus, given that a student must complete a set number of activities to complete a lesson, this measure provided some insight into both a student's progress within the DreamBox curriculum and their level of performance. The median number of lessons students completed in first grade was slightly more than in kindergarten (see Table 3). In first grade, the median student completed 133 lessons (range 38 to 323) compared to 115 (range 38 to 216) for kindergarten (or about 15% more).

Relationship between use and prior ability. The amount of progress students made in DreamBox (i.e., lessons completed) varied significantly by students' prior achievement scores. To explore how both intensity of use and progress varied by prior academic preparation, we created two groups of students based on their prior scores on the MAP math assessment. For kindergarten students, the measure of prior achievement was their fall 2011 math MAP scores. For first grade students, we used spring 2011 math MAP scores (fall scores were only available for first graders who were new students to the school). We identified the median value for the prior MAP scores within kindergarten and first grade and designated a student as being above or below the median value within his or her grade. We then compared the groups on the number of minutes that a student logged as well as the unique number of lessons completed.

Table 4 shows the mean number of minutes and lessons completed for students above and below the median on measures of prior achievement across kindergarten and first grade. Differences in the average number of minutes logged and lessons completed between the two groups with different levels of prior achievement were statistically significant for first grade, but not for kindergarten. Students in first grade with higher prior achievement logged approximately 20% more time on DreamBox (2,346 minutes compared to 1,976 for the median student in each group) and completed about 30% more lessons (155 compared to 117 lessons completed for the median student in each group).

TABLE 4: DIFFERENCES IN INTENSITY OF USE AND LESSONS COMPLETED BY PRIOR ACHIEVEMENT

	STUDENTS BELOW MEDIAN PRIOR ACHIEVEMENT			STUDENTS ABOVE MEDIAN PRIOR ACHIEVEMENT		
	N	MEAN	STANDARD DEVIATION	N	MEAN	STANDARD DEVIATION
MINUTES (INTENSITY OF USE)						
Kindergarten	49	1,875	289	55	1,924	280
Grade 1*	48	1,976	687	57	2,346	963
LESSONS COMPLETED (PROGRESS)						
Kindergarten	49	112	35	55	122	38
Grade 1***	48	117	49	57	155	58

*Note: * = $p < .05$, *** = $p < .001$*

These differences in time-on-task and progress made within DreamBox for students in first grade with different levels of prior achievement are expected given the self-paced, personalized instructional environment adopted by the school. Students who needed more time to complete lessons and master the content could do so, while more advanced students could demonstrate mastery in different content areas perhaps more quickly and learn and practice new topics that were units or grade levels ahead of the rest of the class. In addition, first-grade teachers within KEA had some discretion over how they used online instruction in their classroom, and some teachers may have assigned some students less time on DreamBox and spent more time working directly with these students while others worked on DreamBox.

Impact Analyses

IMPACT DESIGN OVERVIEW

A quasi-experimental design was used to explore the relationship between adoption of the blended learning model and gains in student learning. The design involved comparing the learning outcomes for students enrolled in KEA in school year 2011–12 to a comparison group of similar students enrolled in similar schools that, ideally, did not adopt a blended learning model. Specifically, the comparison group included students in a "virtual comparison group" (VCG). The Northwest Evaluation Association (NWEA) created the VCG by using scores on NWEA's Measures of Academic Progress (MAP) adaptive assessments for similar students in similar schools who took the same test during the same test window.

The impact analysis was limited to kindergarten students only. Similar analyses were not possible for first grade because of the high level of missing fall MAP test scores for a majority of students. In order to conduct an analysis using the NWEA VCG valid scores are needed for both fall and spring administration of the MAP assessment. Unfortunately, KEA only administered the MAP to students in first grade in the fall if they were new to the school. This resulted in a very small number of first graders with valid fall and spring MAP scores (N = 10). Table A1 (see the Appendix for this profile) shows the baseline descriptive statistics for kindergarten students included in the sample for the impact analysis.

One of the primary limitations of an impact analysis using a VCG design is that the design is unable to isolate the effect of the blended learning model from other aspects of the schools' instructional system, including extended days and instructional time, as well as the academic culture of the school and school leadership. Blended learning is a critical component of KEA's instructional system and thus any differences in academic performance between KEA and other schools are at least partially explained by the adoption of blended learning.[37] However, it is not possible using a VCG design to estimate the relative importance of the blended learning model in comparison to other aspects of KEA's instructional system, school culture, or leadership.

IMPACT FINDINGS

Main findings of virtual comparison group analysis. Kindergarten students at KEA outperformed students in the virtual comparison group on both NWEA MAP math and ELA assessments (see Table 5). For math, over two-thirds of the KEA students met or exceeded the spring MAP

TABLE 5. OVERALL DIFFERENCE BETWEEN KIPP EMPOWER ACADEMY AND VIRTUAL COMPARISON GROUP (VCG) ON SPRING NWEA MAP SCORES

	% STUDENTS MEETING OR EXCEEDING VCG SPRING MAP SCORE	MEAN DIFFERENCE IN SPRING MAP SCORES	STANDARD DEVIATION OF SPRING MAP SCORES	N	EFFECT SIZE	P
MAP Math	69.4%	+4.2	9.7	85	0.43	<.001
MAP ELA	73.5%	+9.2	13.2	83	0.69	<.001

[37] *Some schools in the virtual comparison group may also be using blended learning. Therefore, any effect that is directly related to KEA's blended learning model may underestimate the true effect of blended learning instruction relative to teacher-led instruction only.*

TABLE 6. OVERALL DIFFERENCE BETWEEN KIPP EMPOWER ACADEMY AND VIRTUAL COMPARISON GROUP (VCG) ON SPRING NWEA MAP SCORES, BY GENDER

	% STUDENTS MEETING OR EXCEEDING VCG SPRING MAP SCORE	MEAN DIFFERENCE FROM VCG	SD OF MAP SCORES	N	EFFECT SIZE	P
MAP MATH*						
Female	58.3%	2.4	10.1	48	0.25	0.059
Male	83.8%	6.4	9.4	37	0.66	<.001
MAP ELA*						
Female	79.2%	11.8	12.6	48	0.89	<.001
Male	65.7%	5.6	12.7	35	0.42	0.005

** Male and female effect sizes are statistically significantly different from one another for both ELA and mathematics scores (p < .05).*

scores of students in the VCG (69%), while for reading, almost three-quarters of the KEA students did so (74%). Translated to effect sizes, KEA's students outperformed students in the VCG group in both math (effect size = +.43) and reading (effect size = +.69) and both of these differences were statistically significant.

Subgroup analysis. We also found evidence that performance of kindergarten students relative to the VCG group varied by gender but not prior achievement levels (see Table 6). Specifically, male students significantly outperformed female students relative to the VCG group in math (84% compared to 58%). For ELA, however, female students significantly outperformed male students relative to the VCG group (79% compared to 66%). These effects were also statistically significant.

There is no evidence that differences in MAP scores relative to the VCG varied by prior achievement level (see Table 7). KEA students with lower and higher prior achievement levels outperformed the VCG to a similar degree.

TABLE 7: OVERALL DIFFERENCE BETWEEN KIPP EMPOWER ACADEMY AND VIRTUAL COMPARISON GROUP (VCG) ON SPRING NWEA MAP SCORES, BY PRIOR ACHIEVEMENT

	% STUDENTS MEETING OR EXCEEDING VCG SPRING MAP SCORE	MEAN DIFFERENCE FROM VCG	SD OF MAP SCORES	N	EFFECT SIZE	P
MAP MATH						
Higher Prior	69.6%	3.5	8.5	46	0.36	0.008
Lower Prior	69.2%	4.9	9.5	39	0.51	0.001
MAP ELA						
Higher Prior	68.3%	9.2	13.6	41	0.70	<.001
Lower Prior	78.6%	9.1	12.4	42	0.69	<.001

TABLE 8. RESULTS OF ANALYSIS OF DREAMBOX USE PREDICTING EXPECTED STUDENT ACHIEVEMENT

	LOWER THAN PREDICTED SPRING MAP SCORE			HIGHER THAN PREDICTED SPRING MAP SCORE		
	N	MEAN	STANDARD DEVIATION	N	MEAN	STANDARD DEVIATION
DREAMBOX MINUTES						
Kindergarten	45	1,869	305	59	1,926	267
Grade 1	51	2,176	854	51	2,197	901
DREAMBOX LESSONS COMPLETED						
Kindergarten	45	114	41	59	120	33
Grade 1	51	133	55	51	144	60

EXPLORING THE LINK BETWEEN USE AND STUDENT OUTCOMES

Next we explored the relationship between intensity of use, progress made within DreamBox (lessons completed), and student achievement. For this analysis, we controlled for prior achievement, since prior achievement is the strongest predictor of spring test scores, and as shown above, a student's prior achievement level (above or below the school's grade level median) also predicts how much progress the student made in DreamBox. To control for prior achievement we used the score from a previous administration of the NWEA MAP assessment (fall 2011 for kindergarten and spring 2011 for first grade) to predict the spring 2012 MAP score. We then divided students into two groups, those that scored higher than predicted on the spring MAP assessment and those that scored lower, and compared the two groups on the minutes they worked on DreamBox and the number of lessons that the students successfully completed. Table 8 shows results for this analysis. While students in kindergarten and first grade who scored better than predicted by the prior achievement scores on the spring MAP math assessment tended to spend slightly more time working on DreamBox and made slightly more progress than students who scored lower than predicted, the differences were small and not statistically significant. For example, in kindergarten, students that scored higher than predicted on the spring MAP assessment logged an average of 1,926 minutes compared to 1,869 minutes for students who scored lower than predicted on the spring MAP (or 30 minutes more across the entire school year) and completed approximately 6 more lessons (120 compared to 114 or 5% more). In first grade the differences were also small, favoring students with higher than predicted spring MAP scores. Student in first grade with higher than predicted spring MAP scores logged an average of about 20 more minutes on DreamBox (2,197 minutes compared to 2,176) and completed 11 more lessons (144 compared to 133 or 8% more).

Appendix: KIPP Empower Academy

TABLE A1. BASELINE DESCRIPTIVE STATISTICS: STUDENTS IN THE ANALYTIC SAMPLE

VARIABLE	PERCENT	N
ETHNICITY		
American Indian or Alaskan Native	3%	3
Black or African American	84%	97
Chinese	1%	1
Declined to state/Unknown	1%	1
Hispanic	10%	12
White	2%	2
GENDER		
Female	51%	59
Male	49%	57
ENGLISH LANGUAGE LEARNER STATUS (ELL)		
ELL	8%	9
Non-ELL	92%	105
PARTICIPATION IN FEDERAL REDUCED-PRICE LUNCH PROGRAM (FRPL)		
FRPL	89%	102
Non-FRPL	11%	12
SPECIAL EDUCATION STATUS – INDIVIDUALIZED EDUCATION PROGRAM (IEP)		
IEP	6%	7
Non-IEP	94%	109

Teacher Survey Results - KIPP LA

SURVEY ADMINISTRATION

Overall, 7 of 8 KIPP LA teachers implementing blended learning completed the Teacher Survey for an 88% response rate. Table S1 describes the characteristics of the respondents in terms of the subject area(s) they teach and years of teaching experience.

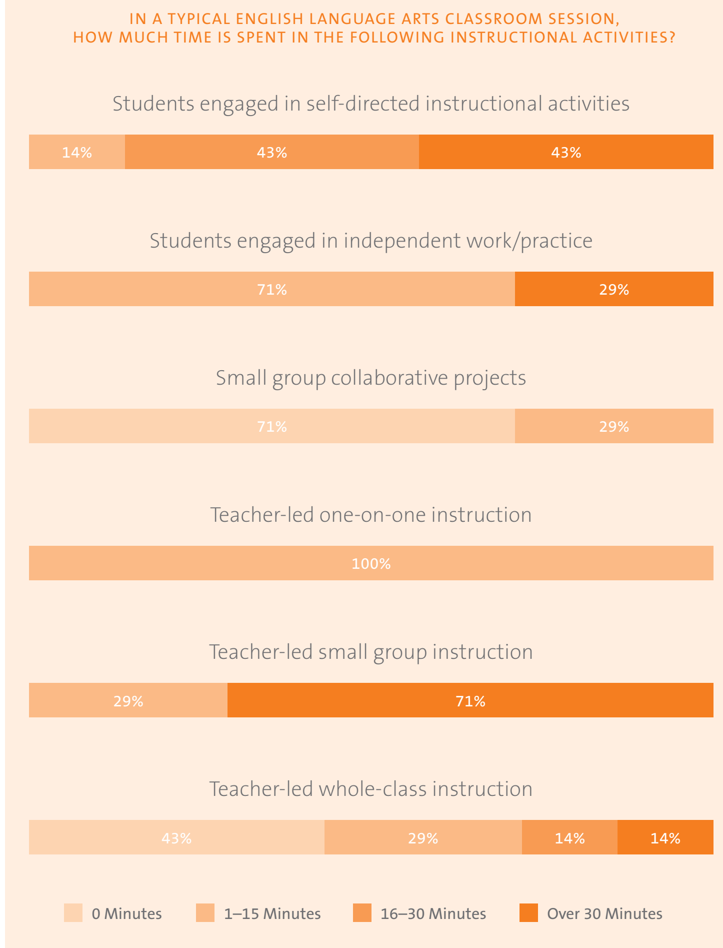
TABLE S1. TEACHER SURVEY RESPONDENT CHARACTERISTICS

TEACHER CHARACTERISTICS	NUMBER	PERCENT
SUBJECT TAUGHT		
English Language Arts (ELA)	1	14%
Math	0	0%
Math and ELA	6	86%
TEACHING EXPERIENCE		
New teachers (3 years or less)	2	29%
Veteran teachers (4 or more years)	5	71%

CLASSROOM INSTRUCTIONAL ACTIVITIES

Teachers reported devoting relatively large amounts of time to small-group instruction, independent work, and self-directed activities during math and English language arts (ELA) classroom sessions and relatively little time in whole-class instruction. In addition, all teachers found time to work with students one-on-one on a daily basis.

FIGURE S1. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (ELA CLASSROOMS)



ENGLISH LANGUAGE ARTS

Most of the KIPP LA teachers surveyed (86%) indicated that students spent at least 16 minutes or more in self-directed instructional activities during a typical ELA class session. Three-quarters (71%) of the teachers indicated that they spent over 30 minutes of a typical ELA classroom session leading small-group instruction, and all (100%) teachers reported spending at least some time providing one-on-one instruction. Relatively less classroom time was typically devoted to small-group collaborative projects and teacher-led whole-class instruction, with 71% and 43% never incorporating these activities, respectively (see Figure S1).

MATH

Most of the KIPP LA teachers surveyed (83%) indicated that a typical math classroom session included over 30 minutes of teacher-led small-group instruction. All (100%) teachers reported spending at least some time providing one-on-one instruction. Relatively less classroom time was typically devoted to small-group collaborative projects and teacher-led whole-class instruction, with 50% and 33% never incorporating these activities, respectively(see Figure S2).

FIGURE S2. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (MATH CLASSROOMS)

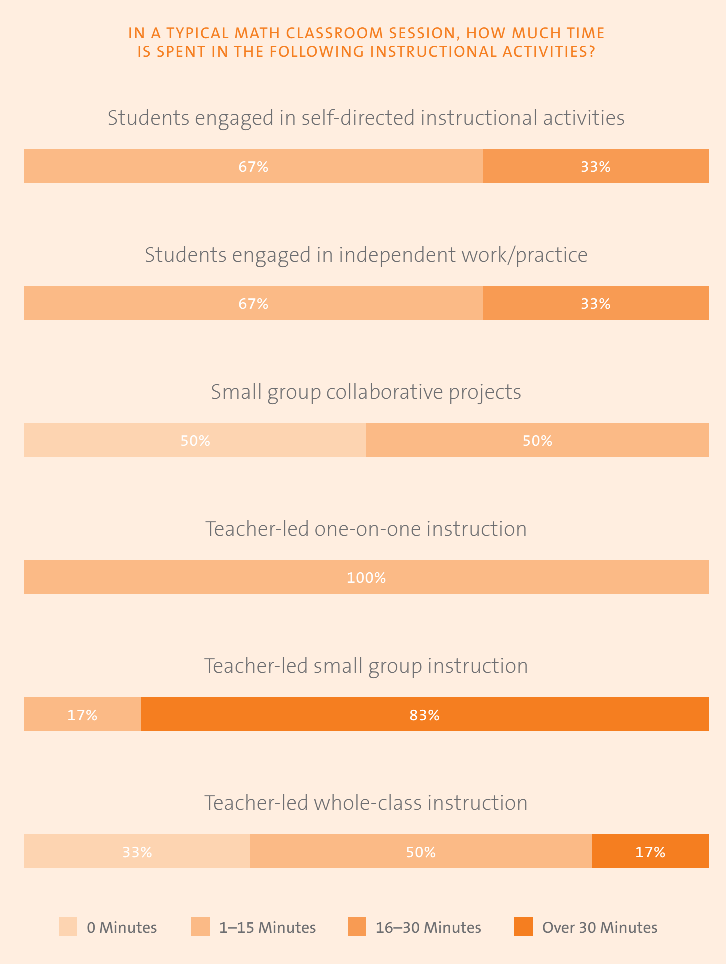
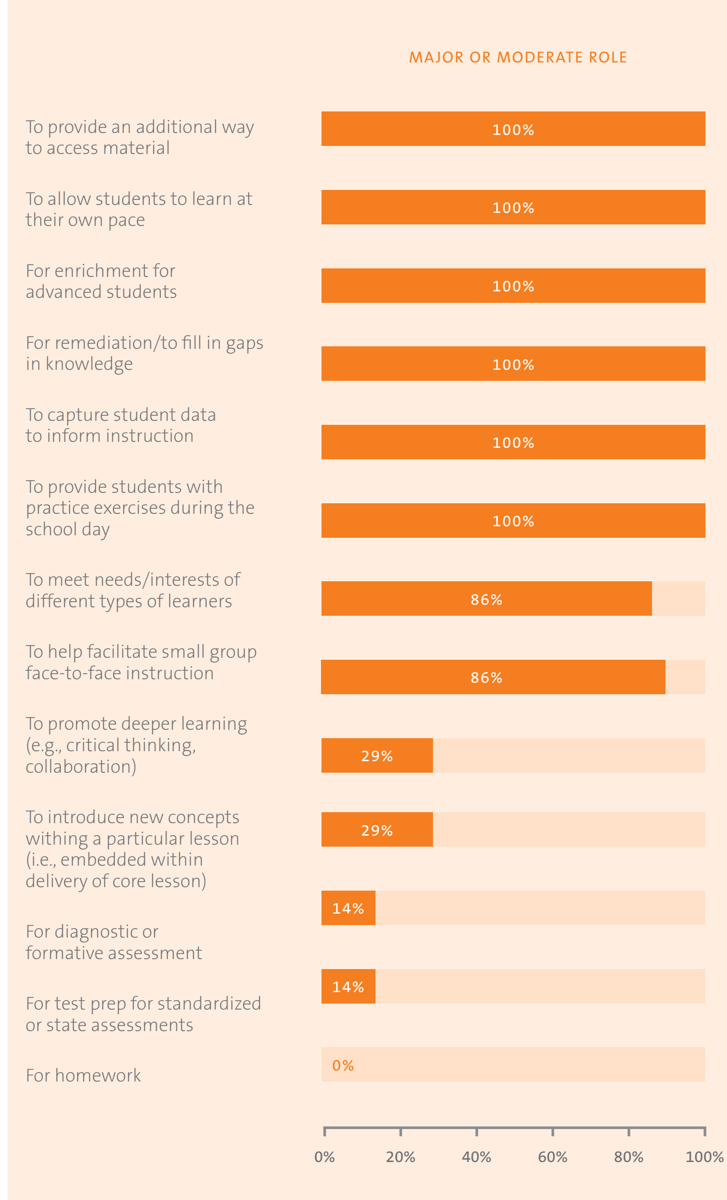


FIGURE S3. PERCENTAGE OF TEACHERS REPORTING DIFFERENT ROLES FOR TECHNOLOGY AND WEB-BASED INSTRUCTION IN SUPPORTING TEACHER-PROVIDED INSTRUCTION



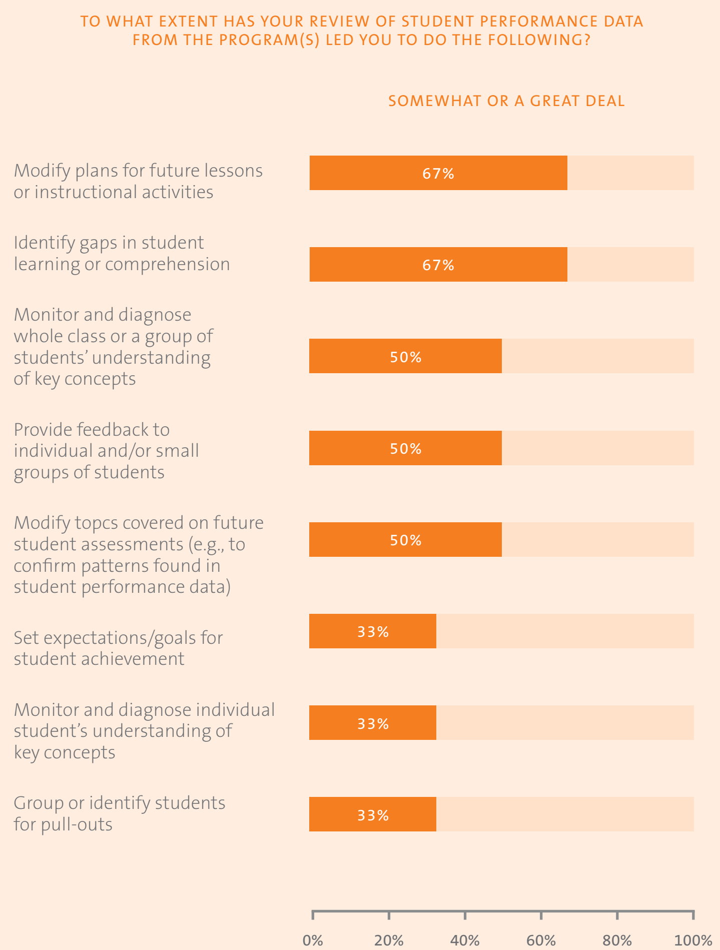
ROLE OF TECHNOLOGY

All KIPP LA teachers indicated that technology and web-based instruction played a “major” or “moderate” role in supporting the instruction they provided by allowing students to learn at their own pace, filling in gaps in student knowledge, offering enrichment for advanced students, providing students with practice exercises, and capturing data to inform instruction (see Figure S3). Many (86%) also noted the role it played in helping to facilitate small-group face-to-face instruction.

USE OF DATA

Most (86%) of the teachers reviewed the student performance data recorded by the software systems used to support the school’s blended learning model; however, the majority of those that did reviewed the data about once a month or less (teachers are provided with reports by the KIPP regional office once a month).

FIGURE S4. PERCENTAGE OF TEACHERS USING DATA “SOMEWHAT” OR “A GREAT DEAL” FOR DIFFERENT INSTRUCTIONAL PURPOSES.



Most teachers found the reports to be “somewhat” or “slightly” useful for informing their instruction, but often relied more on information outside the systems. In general, teachers’ review of the student performance data led them to make decisions or take action at the whole-class level, such as modifying future lesson plans, monitoring and diagnosing whole-class or a group of students’ understanding, and modifying topics covered on future assessments—more so than to diagnose individual students who may be struggling and need additional support (see Figure S4).

TABLE S1: PERCENTAGE OF TEACHERS REPORTING THEY “AGREE” OR “STRONGLY AGREE” THAT BLENDED LEARNING BENEFITS THEIR STUDENTS IN DIFFERENT WAYS

PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS.	PERCENT AGREE OR STRONGLY AGREE
The station model meets the learning needs of my students.	100
Students are highly engaged while using the station model.	100
The station model helps students take ownership for their own learning.	71
Students’ learning and understanding of the materials has improved due to the use of the station model.	71
I would recommend the use of the station model to other teachers.	100

Satisfaction with Blended Learning and Impact on Student Learning

TEACHER SATISFACTION WITH BLENDED LEARNING

The KIPP LA teachers were generally positive in their responses about their school’s blended learning model, with all agreeing that they would recommend the use of the station model to other teachers (see Table S1).

EFFECTIVENESS FOR STUDENT TYPES

Teachers reported the station model was at least somewhat effective for various groups of students, though most effective for students whose academic work was ahead of most students their age (see Figure S5).

EFFECTIVENESS RELATIVE TO NONBLENDED ENVIRONMENT

While half the teachers felt the station model was more effective in impacting a variety of academic skills relative to a nonblended classroom, three of the seven teachers responding had no opinion regarding the

FIGURE S5: PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF EFFECTIVENESS OF THE BLENDED LEARNING MODEL FOR STUDENTS WITH DIFFERENT LEVELS OF ACADEMIC PERFORMANCE OR NEEDS.

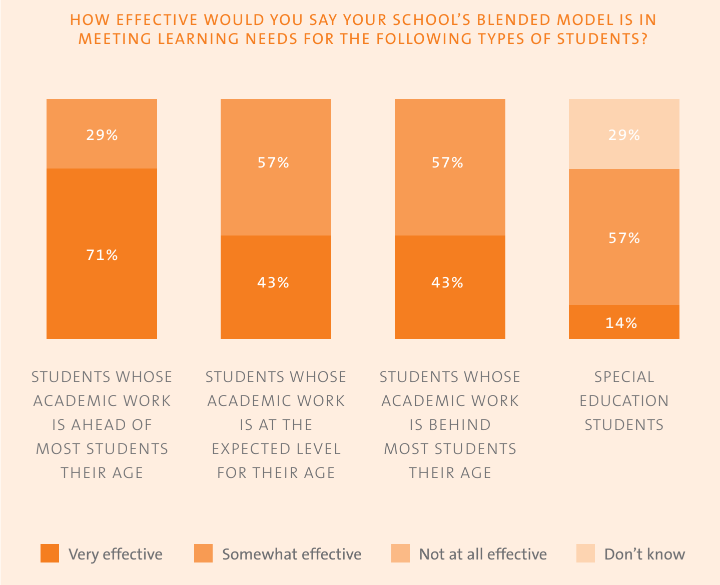
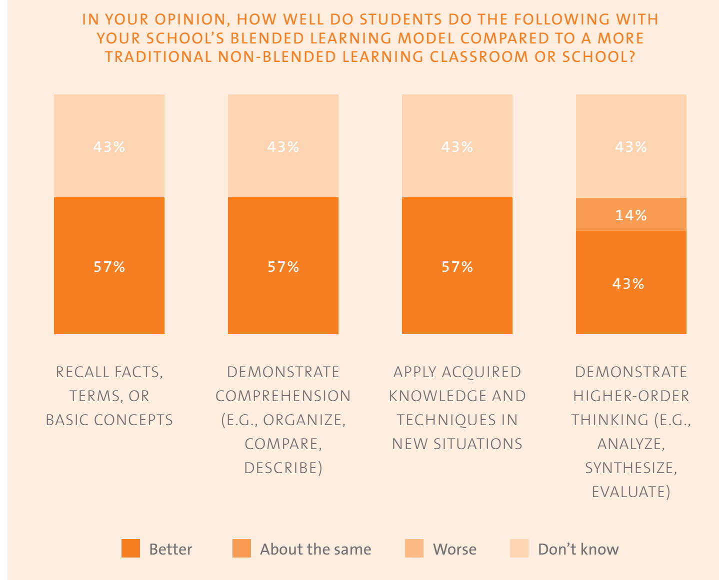


FIGURE S6: PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF THE EXTENT TO WHICH THEIR STUDENTS DEMONSTRATE HIGH-ORDER THINKING AS A RESULT OF THE BLENDED LEARNING MODEL RELATIVE TO A MORE CONVENTIONAL INSTRUCTION MODEL.



impact of the model on students' skills (see Figure S6). Two of seven KIPP LA teachers completing the survey had 3 years of experience or less; both of these teachers selected "don't know" when asked to compare students' skills with the school's blended model to a nonblended environment. In most cases, the veteran teachers with 4 or more years of teaching experience were positive in their ratings, indicating that students did better in these skill areas with blended learning.

TRAINING AND SUPPORT

All but one (71%) of the teachers at KIPP LA participated in a training or orientation session (either in person or online) directly related to their school's use of blended learning or on the specific software program(s) supporting it, and almost all were satisfied with the training they received. Teachers reported spending very little of their own time outside of any professional development (5 hours or less) getting acquainted with the software program(s) or for planning how to best integrate the station model with their instruction.

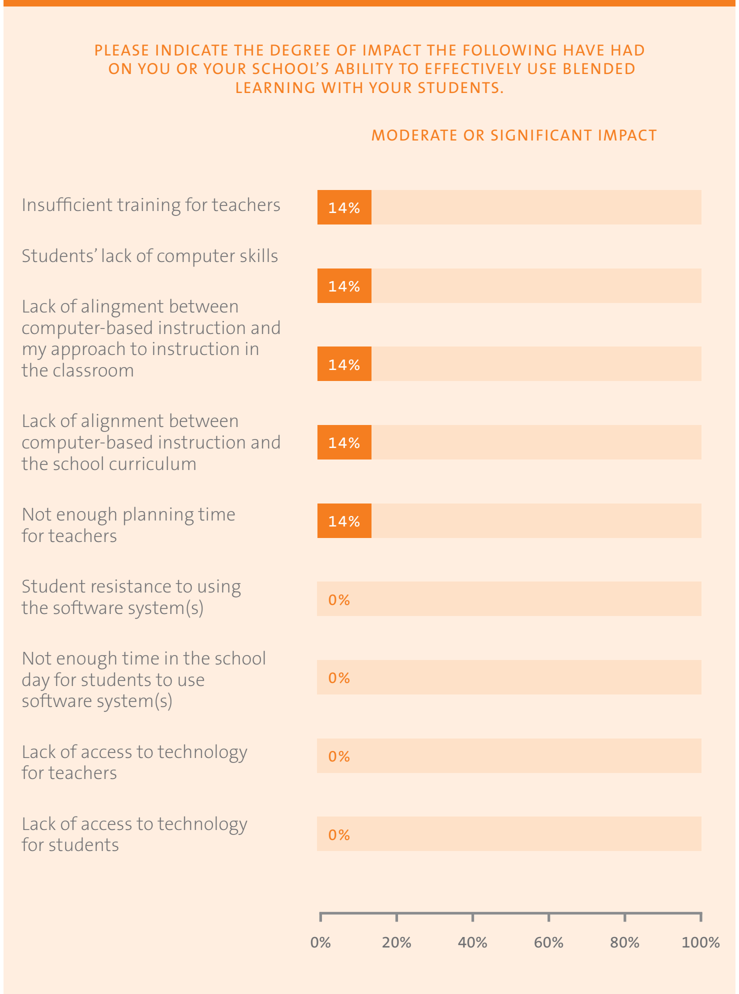
FACTORS INFLUENCING USE

KIPP LA teachers completing the survey reported very few nontechnology barriers to the effective use of blended learning (see Figure S7); almost none of the teachers reported any of the potential challenges listed as having a "moderate" or "significant" degree of impact.

Although the majority (86%) of teachers indicated they or their students had experienced technical problems, none of the teachers indicated they were any more than a "slight" barrier to the effective use of blended learning in the school. The most commonly cited technical problems included software glitches or bugs (83%), Internet connection unreliable/down (67%), insufficient bandwidth/Internet too slow (67%), and student computers not working properly (67%).

Michael & Susan Dell FOUNDATION

FIGURE S7: PERCENTAGE OF TEACHERS REPORTING THE EXTENT TO WHICH DIFFERENT FACTORS INFLUENCED EFFECTIVE USE OF THE BLENDED LEARNING MODEL TO A “MODERATE” OR “SIGNIFICANT” DEGREE



Michael & Susan Dell FOUNDATION

Research Profile: Rocketship Education (2011–12 School Year)

Introduction

Rocketship Education (Rocketship), founded in 2006, is a charter management organization (CMO) that included seven elementary schools serving approximately 3,800 K–5 students in San Jose, California in the 2012-13 school year. Two additional new schools—one in San Jose and one in Milwaukee—opened in fall 2013. The mission of Rocketship is to close the achievement gap by opening high-quality schools in low-income neighborhoods where access to excellent schools is limited. All schools in the Rocketship network implement a rotation blended learning model with online instruction conducted within a learning lab.

Rocketship began implementing their blended learning instructional model at one school during the 2007–08 school year, and eventually expanded the model to a total of five elementary schools by the 2011–12 school year, the year of this study. The primary motivation of Rocketship's blended learning model is to leverage the talent of high-quality teachers and ensure that each student's specific needs are met through a blend of traditional instruction, adaptive technology, and targeted tutoring. Rocketship leverages the talents of classroom teachers by using the learning lab to promote basic-skills development in ELA and math, which frees up teachers in the classroom to focus on the grade-level curriculum and the development and extension of higher-order thinking skills. The use of the learning lab as a key component of the instructional day and staffed by noncertified Individualized Learning Specialists (ILSs) allows schools to offer extended instructional days while operating solely on district, state, and federal tax revenues without the need for ongoing philanthropic support to cover operating expenses.[38]

BLENDED LEARNING MODEL DESCRIPTION

Rocketship had adopted a lab-rotation blended learning model.[39] Each lab held up to 100 students (4 classrooms) supported by 5 adults. Learning labs were monitored by noncertified staff, ILSs, with varying degrees of experience teaching and supporting the learning of elementary-age students.

The ILSs were responsible for managing the transition of 100 students into and out of the learning lab, keeping students on task, reviewing student progress on the online programs, and supporting students if they had a technical problem with hardware or software. While the support ILSs provided students was predominantly nonacademic in nature, two to three ILSs conducted tutoring sessions with small groups of students singled out for needing additional direct instruction during each instructional block. ILSs were present in the learning lab approximately 7 hours per day and they saw every student in the school on a daily basis. Each ILS was responsible for monitoring a class of 20 students within each learning lab, including 5 "focus" students, who needed more intense support based on teacher recommendations and performance on benchmark assessments.

Students participated in two, 30- to 40-minute lab sessions each day, one for math and one for English language arts (ELA).[40] On Fridays students were given free choice over the online programs they chose to work on in the learning labs, including a set of math games available within one of the core online programs. In addition to the learning labs, students also received teacher-led classroom-based instruction in ELA and math. Classroom-based instruction consisted of one daily 200-minute block of literacy instruction (that included social studies and the arts) and one 100-minute block of math instruction (that included instruction in science).

The learning lab was used by Rocketship to supplement face-to-face classroom instruction with the instructional content of the learning lab software programs loosely sequenced with the classroom curriculum. However, given that students' progress in the learning lab was self-paced, the alignment between daily or weekly lessons in the classroom and students' work in the learning lab was limited. Learning lab instruction was used to support basic skills development and to "fill in the gaps," while the primary instruction occurred in the classroom through teacher-led whole-group and small-group instruction leveled by ability.

The learning lab also played a key role in Rocketship's Response to Intervention program (RTI).[41] During the learning lab sessions RTI Tier 2 students were nominated by teachers to receive small-group tutoring delivered by the ILSs.[42] ILSs used a highly scripted curriculum prepared by Rocketship's Individualization Team. The Individualization Team is comprised of curriculum and instructional technology specialists that supply Rocketship schools with the curriculum for the learning lab and make decisions on the selection and sequencing of the online programs. After each benchmark assessment (8-week cycle), the school leadership team reviewed students' and grade-level performances, and decisions were made as to which grade levels and students would be prioritized for the tutoring sessions in each of the schools during the next 8-week cycle.

---

[38] *A detailed description of Rocketship's instructional and operational model is available in a case study developed as part of this research study and published on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).*

[39] *This description captures the rotation model in place for the 2011–12 school year, when this study was conducted. Rocketship is continuously refining its model and approach to blended learning. The learning lab is expected to change and develop in the coming years in order to further improve the level of personalized learning taking place at all Rocketship schools.*

[40] *Part way through the 2011–12 school year, in one or two schools in the Rocketship network, the split-learning lab times for K–1 students were combined into a single 100 minute learning block to reduce the number of daily transitions for the younger students.*

[41] *Response to Intervention programs are a set of practices to help schools continuously identify students with different learning needs and provide interventions for students with the greatest needs. Schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions, and adjust the intensity and nature of those interventions depending on a student's responsiveness. Students are grouped into three tiers based on learning needs and different types of instruction and intervention are provided to each. Tier I students are at or above grade level and do not require intervention; Tier 2 students are 1–3 years behind their peers; and Tier 3 students are 3 or more years behind their peers.*

[42] *RTI Tier 3 students, those with the greatest learning needs, received their teacher-led interventions outside the lab, in a "pull-out" classroom.*

TABLE 1: ROCKETSHIP SCHOOLS IN RESEARCH SAMPLE				
SCHOOL NAME (YEAR OPENED)	GRADE LEVELS 2011–12	NUMBER OF STUDENTS	2011–12 API SCORE[A]	2010–11 SIMILAR SCHOOL RANKING[B]
Mateo Sheedy Elementary (2007)	K–5	K–5	K–5	K–5
Si Se Puede Academy (2009)	K–5	K–5	K–5	K–5
Los Suenos Academy (2010)	K–4	K–4	K–4	K–4
Mosaic Elementary (2011)	K–3	K–3	K–3	K–3
Discovery Prep (2011)	K–4	K–4	K–4	K–4

[A] *The Academic Performance Index (API) ranges between 200 and 1,000, with 800 as the statewide goal for all schools.*

[B] *To create the similar-schools ranking the California Department of Education compares a school's test scores to 100 schools across the state with similar demographic profiles. California uses parent education level, poverty level, student mobility, student ethnicity, and other data to identify similar schools. Rank 1 means the school performed below at least 90 of its 100 similar schools. Rank 10 means the school performed above at least 90 of its 100 similar schools.*

Sample

Table 1 summarizes the study sample for the five Rocketship schools that implemented the rotation learning lab model during the 2011–12 school year: Mosaic Elementary (ROMO), Si Se Puede Academy (RSSP), Los Suenos Academy (RLS), Mateo Sheedy Elementary (RMS), and Discovery Prep (RDP). In the school year 2011–12, two schools enrolled students in Grades K–5 (RSSP and RMS); two schools K–4 (RLS and RDP), and one school K–3 (ROMO). The 2011–12 school year was their first year of operation for RDP and ROMO.

All five Rocketship schools were considered high performing based on their performance on the California Standards Test CST. Academic Performance Index (API) scores for Rocketship schools based on spring 2012 CST scores ranged from 793 to 924, while 2010–11 similar school rankings for RMS, RLS, and RSSP placed these schools in the top 10% of schools serving similar school populations. The historic similar-schools ranking provide an indication of these schools' baseline performance in the 2010–11 school year. Similar-schools rankings were not available for ROMO and RDP, since these schools opened in fall 2011.

TABLE 2. DATA COLLECTION SUMMARY, FALL 2011 AND SPRING 2012 SITE VISITS, ROCKETSHIP		
	FALL 2011 SITE VISITS	SPRING 2012 SITE VISITS
OBSERVATIONS		
Schools	1[A]	3[B]
Classrooms	-	7
Labs	1	3
INTERVIEWS		
Leaders	2	6
Teachers	-	5
Students	-	-
Lab Facilitators	-	4

[A] *Mateo Sheedy Elementary*

[B] *Mosaic, Los Suenos, Si Se Puede*

Michael & Susan Dell FOUNDATION

TABLE 3. RESPONSE RATES, TEACHER AND LAB FACILITATOR SURVEY, ROCKETSHIP

INSTRUMENT	RESPONSE RATE
Teacher Survey	73% (51 of 70)
Lab Facilitator Survey	86% (24 of 28)

TABLE 4. TEACHER CHARACTERISTICS, TEACHER SURVEY, ROCKETSHIP

TEACHER CHARACTERISTICS	NUMBER	PERCENT
SUBJECT TAUGHT		
English Language Arts (ELA)	34	67
Math	17	33
Math and ELA	1	2
TEACHING EXPERIENCE		
New teachers (3 years or less)	29	57
Veteran teachers (4 or more years)	22	43

Data Collection

SITE VISITS

Two members of the research team conducted one-day site visits to select Rocketship schools in fall 2011 and again in spring 2012. Table 2 summarizes the number of observation and interviews conducted during each site visit. For the fall 2011 site visit, the research team only conducted observations and interviews within RMS. For the spring 2012 site visits, the research team traveled to ROMO, RLS, and RSSP.

TEACHER SURVEY

Within Rocketship schools, SRI conducted a survey of teachers and lab facilitators implementing blended learning. Table 3 shows the response rate for each survey, with 73% of Rocketship teachers responding (51 of 70) and 86% of ILSs responding (24 of 28). Table 4 shows the characteristics of the teacher respondents in terms of the subject areas they teach and years of teaching experience.

SYSTEM LOG DATA

We collected and analyzed system log data to gain a better understanding of the extent to which students used the online programs in the learning lab across all five Rocketship elementary schools. The analysis was constrained to Grades K–3, the grades common to all Rocketship schools during the 2011–12 school year.[43] The analysis was also limited to online math instruction and included students' use of the most frequently used math online programs: DreamBox, ST Math, and TenMarks.[44] For ELA instruction that took place in the learning lab during the 2011–12 school year, Rocketship used Accelerated Reader, a leveled reading program where students read print-based books appropriate for their reading level and then completed reading comprehension assessments on the computers. Since the students did not receive ELA instruction online, students' use of Accelerated Reader was not included in the system log analysis. In addition, schools also used Compass Learning in the learning lab to support ELA instruction. According to Rocketship leadership, it was used most consistently by schools in kindergarten than other grades. The original intent was to split the lab time between Accelerated Reader and Compass Learning, but few schools did that, with most prioritizing Accelerated Reader in upper grades and only using Compass Learning a few days a week or for limited time, if at all. As a result, use data from Compass Learning was not accessed by the research team and was not included in these analyses.

To measure total intensity of use for the DreamBox system, the amount of time students spent on each DreamBox "session" was summed across the entire school year. A session is defined by a student logging into DreamBox and subsequently logging out. For

[43] *Grade 3 was the highest grade common to all schools. During the 2011–12 school year, one school enrolled students in K–3 (ROMO), two schools had grades K–4 (RDP and RLS), and two schools Grades K–5 (RSSP and RMS).*

[44] *During the 2012–13 school year there were several changes to the mix of online programs used in the learning lab to support math instruction. Curriculum Associates' iReady math was added, while TenMarks was dropped from the online program offerings.*

ST Math, the number of minutes that a student logged per "syllabus objective" was summed for the entire year. For TenMarks, like DreamBox, minutes logged for each session were totaled for the academic year.[45, 46]

STUDENT LEARNING OUTCOMES

The learning outcomes analyzed were individual student scores on the Northwest Evaluation Association's (NWEA) Measurement of Academic Progress (MAP) assessments in reading and math. The NWEA MAP is a computer-based adaptive assessment that is aligned with the California Common Core State Standards (CCSS) (http://www.nwea.org). Rocketship schools administer the NWEA MAP assessment three times per year (fall, winter, and spring). NWEA MAP scores were provided by Rocketship, and NWEA provided MAP scores for a "virtual comparison group" (VCG) created for the research team using NWEA MAP scores for similar students in similar schools who took the MAP during the same test window.

For the achievement gap analysis (described below), scores on the CST were analyzed to examine the extent to which Rocketship schools were able to reduce the high-income/low-income achievement gap in the state.

Implementation Findings

Findings from the site visits and teacher survey were used to help understand the facilitating factors and barriers to implementing the rotation blended learning model at Rocketship during the 2011–12 school year. Additionally, an analysis of the system log data from the primary online programs used in the learning lab to support math instruction was used to help better understand the intensity with which students used the online systems to support math and the potential connection between use and learning.

FINDINGS FROM THE SITE VISITS AND TEACHER SURVEY

Data from the fall 2011 and spring 2012 site visits and the teacher and ILS surveys were used to report on different aspects of implementation of Rocketship's rotation model. These findings are discussed below by theme, including infrastructure and technology issues, the role of the learning lab in instruction, the support and training teachers received, and teachers' use of the system data to support their instruction. In addition, we report on teachers' overall satisfaction with the blended learning model and efforts to improve student engagement. A summary of the survey results is included in the Appendix for this profile.

INFRASTRUCTURE AND TECHNOLOGY

One-to-one computing and a single sign-on portal supported the implementation of the learning lab model. Rocketship implemented a blended model in which students engaged in online instruction in math and ELA in a learning lab. During the 2011–12 these labs had enough computers available at all times to support one-to-one computing for up to 100 students (4 classrooms). Each student was seated at their own learning carrel with cardboard or wood dividers on either side to prevent interactions between students. Students accessed the online programs through a single sign-on portal that allowed very young children to log in and out of the various online programs efficiently with little or no help from adults.

Technical challenges impacted students' use of online instruction. Nearly all (92%) ILSs reported that they or their students experienced some sort of technical problem during the school year within the learning lab, and of these ILSs, approximately half reported the technical problems were a moderate or significant barrier that impacted the effectiveness of the learning lab instruction. According to assistant principals and the ILSs interviewed, the technical problems mostly occurred at the beginning of the year and included

[45] *TenMarks, however, provided a "multiple sessions" output value in place of a numeric value when a student worked on content over multiple log-ins and log-outs. To include these sessions in our time estimates, we imputed "multiple session" values by identifying the average number of minutes students spent per lesson and multiplied it by 1.5. These imputed values were added to nonimputed values to create the total amount of time students spent on TenMarks.*

[46] *The bulk of time students worked on the TenMarks program was spent solving multiple-choice math problems generated by the program. Other math activities within TenMarks included a series of math games that students were allowed to access on Fridays, their "free choice" day. Data related to the time students logged on to these math games was not available and thus was not included in the system-log data analysis. Intensity of use measures for each system were then merged with students' demographic information and assessment data and then analyzed.*

unreliable Internet connectivity, intermittent online program "crashes," and frequent problems with the single sign-on portal that prevented students from accessing the software. While problems with Internet connectivity were resolved in most schools early in the school year, problems with the sign-on portal continued throughout the year.

THE ROLE OF THE LEARNING LAB IN INSTRUCTION

Connection between the classroom and the learning lab was limited. The instruction received by students in the learning lab was designed to supplement the instruction in the classroom by providing students with opportunities to practice skills and to fill-in gaps in foundational knowledge. The most frequently used online programs used to support math instruction—DreamBox and ST Math—were adaptive, allowing students to work at their own level and pace. During the 2011–12 school year there was a limited attempt by Rocketship to align the sequence and pacing of the online math instruction with the classroom instruction. Rocketship worked with vendors to varying degrees to map their classroom lessons to units in the online programs. They then shared their pacing guides for classroom instruction with vendors so that, in theory, when students were in the lab they were working on content that was being taught around the same time in the classroom. However, since students worked at their own pace in the learning lab on DreamBox and ST Math, on any given day math teachers could be covering one set of learning objectives in the classroom with all students while these same students in the learning lab might be working on a different set of skills that were behind or ahead of the classroom curriculum. In our interviews with teachers from one school, the math teachers talked about the desire for improved integration between the classroom and learning lab through providing teachers with the ability to assign the content of the online instruction based on students' performance in the classroom. Of the three programs used in the learning lab to support math instruction, ST Math and TenMarks are assignable (ST Math to a limited degree), while DreamBox is not.

The sense of separation between the classroom and the learning lab instruction was enhanced due to the lack of day-to-day involvement of teachers in the learning lab. During the 2011–12 school year, teachers' physical presence in the learning labs was limited, with 65% of teachers reporting on the survey that they never supervised students in the learning lab. Formal opportunities for teachers to interact with the ILSs to share information about students were also limited. About half of teachers (46%) reported that they met with the ILSs to discuss student performance in the classroom and learning lab once a month or less, primarily during "data days" when students' performance on benchmark assessments were shared with teachers and ILSs (every 8 weeks for math and ELA). Both teachers and ILSs reported on the surveys that they wanted to meet more frequently and that it is important that they know what is going on, particularly what students struggle with, in both the classroom (for the ILSs) and the learning lab (for teachers) to be effective in their respective roles.

Reports from assistant principals and teachers within schools visited in the spring suggest that schools were working on ways to facilitate a greater connection between the classroom and learning labs and the sharing of information about students between teachers and ILSs.[47] For example, one school had arranged common lunch times for teachers and ILSs so they could interact more frequently. This same school, during the second semester, allowed teachers to spend 20 minutes of one class period every two weeks observing in the learning lab while ILS was assigned to cover their classroom. In addition, in some schools ILSs were working with teachers to extend the scripted curriculum used during the tutorial sessions in the learning lab with materials from the classroom.

TEACHER TRAINING AND SUPPORT

While most teachers received some training related to their school's blended learning model, a majority reported it was not adequate for their needs. Most (86%) Rocketship teachers reported they had participated in some (30%) or all (56%) of the training sessions provided directly related to their school's use of blended learning or on the specific software program(s) supporting it. However, among those who did participate, only 39% were satisfied with the training they received, while 61% indicated they were dissatisfied. A number of the teachers who indicated they were not satisfied with the training also commented that they lacked sufficient information about what their students were doing in the lab (e.g., what content they were exposed to, how much time they were spending on different programs). Teachers also cited either a lack of data or access to data from the online instructional programs, as well as not being provided

[47] *The need for more integration between classroom and online learning is influencing how Rocketship evolves its model to best serve students.*

with enough information about how to use the online data to assess student progress and inform their instruction. More than 90% of teachers responded that it was "very important" to receive training on how to access and interpret the student progress and performance data available from the online programs, and almost three-quarters of teachers (73%) reported that it was very important to receive training on how to use the data to inform their instruction.

The role of centralized support from Rocketship's National Team was critical to the implementation of the CMO's blended learning model. Curriculum and technology specialists, operating out of Rocketship's national office, were responsible for the design of the learning lab and its curriculum, as well as the selection and ongoing evaluation of the online programs used in the lab. They were also responsible for the aggregation and reporting of student performance data to schools and teachers throughout the school year on a series of common and regular benchmark assessments administered by the schools in ELA and math. The National Team was continuously soliciting feedback from schools and teachers regarding the implementation of the learning labs and designing and piloting refinements of the model during the school year.

The curriculum and technology specialists also established criteria for the selection of online software and continue to monitor the market for online programs that may improve the instructional experience for students and the feedback provided to teachers on student performance. The six criteria include content that is aligned to the Common Core State Standards; an application program interface (API) that supports single sign on, management of student accounts and data integration; adaptivity; assignability; quality assessments; and affordability.

Assistant principals played a critical role in managing the learning labs and coaching the ILSs. Assistant principals (APs) worked with their ILS team early in the school year to set expectations for operations of the learning lab, including efficiency of the transitions, overall academic culture of the learning lab (e.g., "college library quiet"), norms for student behaviors, and processes for monitoring student engagement and providing students with timely support. Each week the AP and ILSs set goals for student academic progress while in the learning lab, including the number of quizzes students should pass in Accelerated Reader or the number of lessons or units completed in a math online program. Each AP met with their school's ILS team every Friday from 2:30 p.m. to 4:00 p.m. to review system log data against the weekly goals; in addition, some APs also had weekly one-on-one meetings with ILSs. During the weekly meeting the online data was disaggregated for each group of students assigned to an ILS, and at least in one school, for a set of five "focus" students that had been identified as needing additional support. At the end of the meeting, the AP and ILSs agreed on one "focus area" for the following week to improve the effectiveness of students' time in the learning lab (e.g., making the transitions more orderly and efficient or initiating and tracking check-ins with each focus student). The APs were regularly in the learning lab observing and could provide real-time coaching to the ILSs, particularly in the beginning of the year and when working with newly hired, less-experienced ILSs. A wireless communications system was used, with the ILSs wearing ear buds, which allowed the APs to monitor the lab, paying particular attention to that week's focus practice or procedure, and then to communicate with individual ILSs as needed, providing real-time feedback and encouragement.

KEY FACTORS SUPPORTING BLENDED LEARNING ADOPTION

Nontechnology barriers to effective use varied for teachers and ILSs. Rocketship teachers and lab monitors reported different sets of barriers that impacted their school's ability to effectively implement the rotational blended learning lab. A majority of teachers reported that not enough planning time (88%), lack of alignment between the instruction in the learning lab and school curriculum (71%), insufficient training (69%), and the lack of alignment between the online and their own approaches to instruction all had moderate or significant impacts on their schools' ability to effectively implement the blended learning model. In contrast, a majority of lab monitors reported that a lack of student engagement or motivation and student resistance to using the online programs were the factors that had moderate or significant impacts on their school's ability to effectively use blended learning with their students.

Daily learning logs were implemented to improve student engagement. The ILSs interviewed in one school reported that they felt that many students did not consider the learning lab as an extension of the classroom, primarily due to the absence of their classroom teacher in the lab. As a result, according to these ILSs some students were less engaged and more resistant to completing the expected work than they would have been in the regular classroom, particularly older students. Another factor that may have contributed to some students' wavering engagement, according to the APs and ILSs interviewed, was that over the course of the school year, students found some of the online programs more engaging than others. According to those interviewed, some students favored working on particular programs—DreamBox , ST Math, and Equatia—and resisted working on the others, particularly in the latter half of the school year after the novelty of the online programs and the learning lab had started to fade.

To keep students engaged in the learning lab and to increase students' sense of ownership over their work, both teachers and ILSs emphasized during our interviews the importance of teachers communicating to students that what they do in the learning lab matters. During the spring, some schools instituted a learning log that had students record weekly goals for their progress in ELA (e.g., number of Accelerated Readers passed) and math (e.g., how many units they were going to complete in the online programs). Students used the logs to record their daily progress in the programs, as well as to document their work as they worked through solutions to math problems in the online programs. The logs were shared with the classroom teachers each Friday and teachers were encouraged to review the logs and, in the classroom, to publicly celebrate those students who met their weekly goals, thereby sending signals to the students that their work in the learning lab mattered and that their classroom teacher was monitoring their progress. The teachers interviewed were very positive about the use of the learning logs, and math teachers in particular reported that the learning logs provided, for the first time, a view of what their students were learning in the lab.

According to Rocketship leadership and assistant principals, the "culture" established in the learning lab by the ILSs was an important factor determining the level of student productivity within the lab. A lot of emphasis was placed on helping students make efficient transitions into and out of the learning lab. Students were expected to transition in and out of the lab quietly and begin work immediately on the online software programs without interacting with other students. Assistant principals worked with the ILSs to establish procedures to support efficient transitions and maintain student focus within the labs. In addition, each ILS team developed their own systems to maintain a productive learning culture within the lab. For example, in one school, students used nonverbal cues by raising different colored popsicle sticks to notify the ILSs when they needed help with a technical problem, academic support, or a bathroom break.

USE OF DATA TO INFORM INSTRUCTION

Approximately three-quarters (71%) of teachers reported that they reviewed student performance data recorded by the online programs and more than half (53%) reviewed the data at least once a week, while one in three teachers looked at the data about once a month or less. Math teachers reported looking at the student performance data captured by the program on a less frequent basis than ELA teachers. Teachers cited multiple reasons for not reviewing the data at all, or once a month or less, including not having a way to easily access reports of student performance data; relying more on information outside the learning lab (including Rocketship's benchmark assessments as well as teachers' own formative assessments); and not understanding how to use the data from the online programs to inform their instruction.

Almost two-thirds (60%) of teachers reported that the system log data from the online programs was only slightly useful or not at all useful for informing their instruction. Despite the differences in their frequency of review, math and ELA teachers provided similar reports regarding the usefulness of the data for informing their instruction.

At the end of each day, ELA teachers received data on the number of reading comprehension assessments on Accelerated Reader that students passed. Starting in the second semester, math teachers also received data on students' performance on weekly math assessments. The weekly assessments were developed by Rocketship and covered the specific learning objectives (or "micro-objectives") that teachers were to cover in the classroom. Math teachers received data on the number of micro-objectives a student passed at the end of each week. Math teachers did not receive any data related to students' progress and performance within the individual online programs. While math teachers had information on how many learning objectives a student passed, no information was provided on which specific objectives were passed and which objectives students struggled with. In addition, for most of the semester, results were reported at the grade level and not the classroom level. As a result, math teachers in particular found the weekly student progress reports of limited value as a tool to inform their classroom instruction. In addition, in our interviews with assistant principals, teachers, and ILSs we learned that teachers did not trust the micro-objective assessment data that was coming from the learning lab because the results were not consistent with their own observations of individual student performance in the classroom. Students who performed well on a teacher's formative assessment in the classroom did not necessarily do well on the weekly assessments administered in the learning lab. One contributing factor to this discrepancy may have been students' perception of the learning lab as being different from the classroom in terms for expectations for academic performance, primarily due to the absence in the learning lab of their classroom teacher. In our interview with ILSs, they reported that many students, in their observation, did not consider the learning lab as an extension of classroom learning and therefore did not take seriously the weekly assessments delivered online in the learning lab. As a result, for the first time, teachers were planning to be in the learning lab during the spring 2012 online administration of the NWEA MAP assessment to ensure that students understood the importance of doing well on the assessment (during the 2011–12 school year all of Rocketship's own benchmark assessments were administered in the classroom).

SATISFACTION WITH THE BLENDED LEARNING MODEL

Overall, teachers appeared to be only moderately satisfied with the role of Rocketship's blended learning model in instruction. One-half of teachers surveyed (51%) reported they would recommend the blended instructional model used during the 2011–12 school year to other teachers, 18% would not, and almost one in three (31%) reported they "don't know." These reported levels of satisfaction are consistent across math and ELA teachers. The number of teachers responding "don't know" (14 of 45) was likely a sign that teachers felt they did not have enough exposure to the learning lab or access to reliable data on student performance in the lab to judge its effectiveness in improving student learning. However, when teachers were asked whether their schools' blended learning model met the needs of all students, a majority of teachers reported that the rotation learning lab model was either "somewhat" or "very effective" for students with different levels of academic preparation—for students working above grade level (76%), at grade level (80%), and below grade level (64%) with approximately 20% of teachers responding "don't know." More than one in five teachers (22% or 10 of 45 teachers) reported that they believed that the learning lab model was "not at all effective" for students working below grade level. However, this view varied by school and was limited to teachers in three of the five schools. In contrast, only two teachers reported that the learning lab model was likely "not at all effective" for advanced students, and no teachers held this view for students performing at grade level.

FINDINGS FROM ANALYSIS OF SYSTEM LOG DATA FROM ONLINE MATH PROGRAMS

To capture the intensity with which students used the different online math learning systems, we computed the number of minutes that a student logged over the course of the school year for each of the online math programs.[48] The analysis included all students, including those students who spent a portion of their learning lab time receiving direct small-group instruction from an ILS.

As part of the analysis of system log data we also examined how use of the different online programs varied across schools. When examining cross-school variation in use, we limited the analysis to use in Grades 1 and 3 for illustrative purposes.[49] The core online math program used in the learning lab by students in Grade 1 was DreamBox, although one school, RDP, did provide more than half of the first-grade students an opportunity to use ST Math. In grade 3, three math programs were used, TenMarks, ST Math, and Equatia. During the 2011–12 school year, ST Math was used on Mondays and Wednesdays and TenMarks on Tuesdays and Thursdays. Equatia, an "open world, role playing, adventure game," was assigned to and used by students to develop math fluency during the first 10 minutes of the lab time. On Fridays, students got to choose which programs they wanted to work on and many students chose the math games available through TenMarks according to the ILSs interviewed. Data on the amount of time spent on the TenMarks's games was not available and therefore are not included in these analyses. In addition, students' time spent working on Equatia was also not available for analysis.

[48] *For ELA instruction in the learning lab during the 2011–12 school year, Rocketship used Accelerated Reader, a leveled reading program where students read print-based books appropriate for their reading level then complete reading comprehension assessments on the computers. Since the students do not receive ELA instruction online, students' use of Accelerated Reader was not included in the system log analysis. In addition, schools also used Compass Learning in the learning lab to support ELA instruction. Student system log data associated with use of Compass Learning was not accessed by the research team and therefore is not included in these analyses.*

[49] *Grade 3 was the highest grade common to all schools. During the 2011–12 school year, one school enrolled students in K–3 (ROMO), two schools K–4 (RDP and RLS), and two schools K–5 (RSSP and RMS).*

TABLE 5. DESCRIPTION OF INTENSITY OF USE OF ONLINE MATH PROGRAMS BY GRADE FOR ALL SCHOOLS

STANDARD DEVIATION (SD)	N (OVERALL)	MEAN	STANDARD DEVIATION (SD)	MEDIAN	MIN	MAX
Kindergarten	605 (605)	3,224	806	3,201	258	5,322
1	481 (484)	3,237	893	3,201	75	5,394
2	449 (450)	1,976	715	1,888	139	5,969
3	426 (439)	1,483	797	1,373	145	6,402

Overall intensity of use across schools and within schools is described below, along with how time in online instruction varied by students' prior achievement levels. Tables 5–7 provide descriptive statistics for the number of minutes that a student spent on each system broken out by grade level, across all schools (Table 5) and by individual school for Grades 1 and 3 (Tables 6 and 7). Figure 1 shows how the total minutes of online math instruction was distributed across grade levels within each school. Figures 2 (Grade 1) and 3 (Grade 3) show how the use of each of the programs varied by prior academic performance (performance on the fall 2011 NWEA MAP assessment) by school.

Students spent more time in online math instruction in Grades K–1 than in Grades 2 and 3 (Table 5 and Figure 1). Students in kindergarten and first grade spent approximately twice as much time working in online instruction to support their math learning (about 3,200 minutes or 89 minutes per week over a 36 week school year) compared to students in Grades 2 and 3 (2,000 to 1,500 minutes or 55 to 42 minutes per week).

Several factors may have contributed to this difference between grade levels in time-on-task, including differences in the online software packages used between the different grade levels, differences in how online math and ELA instruction was distributed across the year, and differences in the amount of tutoring that was offered within the lab. Students in K–1 used DreamBox exclusively for online math instruction, while students in the other grades used a combination of TenMarks and ST Math. Assistant principals and ILSs reported that younger students really looked forward to their time in the learning lab and engagement was relatively high throughout the year, while behavior and discipline issues were more common with the older students, particularly in the latter half of the school year. In addition, ILSs reported that throughout the year technical problems limited students' access to TenMarks and in general students found DreamBox and ST Math to be more engaging than TenMarks.
In addition, according to Rocketship leadership, kindergarten students worked exclusively on math programs (i.e., Dreambox) for the first half of the year while in the lab and not on ELA programs. Even though all students had a daily "math block" and "ELA block" in the lab, kindergarten students spent both blocks working on math. Starting in January, if these students had advanced to a certain reading level (Developmental Reading Assessment Level 3), then they started using Accelerated Reader during their ELA block. Many of these students did not reach the advanced reading level until March according to Rocketship leaders.[50]

Finally, the difference between grade levels in use of the online programs may have been the result of additional instructor-led tutoring sessions within the learning lab targeting students in Grades 2–5. According to Rocketship leadership, schools often prioritized students in Grades 2–5 for tutoring interventions because these kids were often further behind academically relative to students in K–1 and the need was more urgent. In California, state testing begins in Grade 2.

Time spent in online instruction varied considerably across schools (Tables 6 and 7, Figure 1). First-grade students in RDP and RLS and third-grade students in RLS spent significantly more minutes working with the online instructional programs than students in similar grade levels in the other schools. During the school year first-grade students in RLS and RDP were active within the online math programs for close to 3,700 and 4,100 minutes, respectively (103 to 114 minutes per week) (Table 6), as much as 35% to 50% more than the other schools, and between 70%

[50] *In addition some schools also did not start first graders on Accelerated Reader until a couple of months into the school year, and instead allowed these students to work on DreamBox during the math and ELA blocks in the lab.*

TABLE 6. DESCRIPTION OF INTENSITY OF USE OF ONLINE MATH PROGRAMS BY SCHOOL, GRADE 1

	DREAMBOX		ST MATH		OVERALL	
SCHOOL[A]	N	MEDIAN MINUTES (MIN/MAX)	N	MEDIAN MINUTES (MIN/MAX)	N	MEDIAN (MEAN)
RDP	84	3,851 (2,324/4,497)	56	618 (125/967)	85 (85)	4,119 (4,148)
RLS	102	3,724 (1,754/5,041)			102 (102)	3,724 (3,642)
RMS	94	2,664 (243/3,553)			94 (95)	2,664 (2,597)
ROMO	99	3,020 (75/5,008)			99 (101)	3,020 (3,003)
RSSP	101	3,046 (1,188/4,160)			101 (101)	3,046 (2,887)

[A] *RDP = Discovery Prep, RLS = Los Suenos Academy, RMS = Mateo Sheedy Elementary ROMO = Mosaic Elementary, RSSP = Si Se Puede Academy*

TABLE 7. DESCRIPTION OF INTENSITY OF USE OF ONLINE MATH PROGRAMS BY SCHOOL, GRADE 3

	TENMARKS		ST MATH		OVERALL	
SCHOOL[A]	N	MEDIAN MINUTES (MIN/MAX)	N	MEDIAN MINUTES (MIN/MAX)	N	MEDIAN (MEAN)
RDP	45	270 (4/683)	70	766 (106/6,183)	71 (75)	1,057 (1,230)
RLS	55	564 (25/912)	93	1,764 (313/5,213)	95 (98)	2,069 (2,155)
RMS	59	481 (41/951)	91	993 (220/2,155)	91 (91)	1,293 (1,363)
ROMO	48	359 (62/905)	68	1,104 (268/2,827)	70 (73)	1,182 (1,396)
RSSP	60	340 (37/717)	98	950 (226/2,200)	99 (102)	1,142 1(1,194)

[A] *RDP = Discovery Prep, RLS = Los Suenos Academy, RMS = Mateo Sheedy Elementary ROMO = Mosaic Elementary, RSSP = Si Se Puede Academy*

to 85% of the 120 to 150 minutes per week allocated in the learning lab for math instruction. In third grade, it was students in RLS who logged considerably more time in the online math programs than students in the other schools, as much as 50% more. The average third-grade student in RLS spent about 2,000 minutes in online math instruction (55 minutes per week or less than half the weekly time allocated to online instruction) (Table 7) compared to the students in the other schools whose use of the programs averaged about 30 minutes per week. Differences between schools in the amount of time students were active in online instruction within the learning lab may reflect differences in the productivity of the "learning culture" established within the different school labs. RLS, the school whose students consistently logged the most minutes across Grades 1 and 3, invested a considerable amount of time and effort at the beginning of the second semester to refine the learning lab culture and management systems, including hiring a new assistant principal in January 2012 to lead the reform.

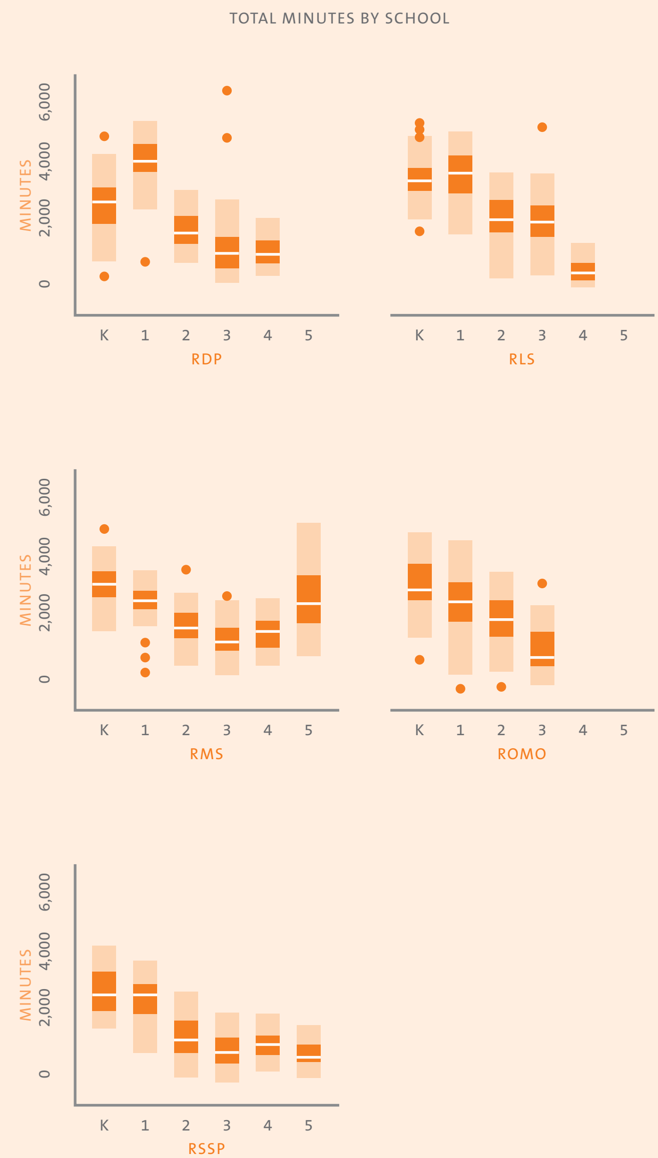
In addition, the differences among the schools in amount of online math instruction experienced by students may have at least partially been attributed to a pilot program that was run across several of the schools in the second semester. According to Rocketship leadership, during the last 8 weeks of school, half of the third and fourth graders at RLS and kindergarten students and first and third graders at RDP spent their ELA learning lab block working on math programs instead. Therefore, many of these students in these grades and schools would have had double blocks of online math instruction in the learning lab during this time.

Students with higher prior academic performance spent more total time in online math instruction than their peers (Figures 2 and 3). On average students with higher prior academic achievement, as measured by performance on the fall NWEA MAP assessment, logged more minutes in online instruction than students with lower performance. This relationship holds for both Grades 1 and 3 but it is clearly stronger for Grade 3. The relationship also holds across most schools—RSSP, ROMO, RMS, RLS (first-grade only) —but not RDP, where in third grade, students with the greatest academic needs logged about 50% more time than others.

Students in third grade in RSSP, ROMO, and RMS with lower academic prior performance—and greatest academic needs—logged about 30% to 60% less time in online instruction than their peers with higher prior academic performance. Students in third grade in these schools who scored at the 20th percentile of fall MAP scores logged about 1,000 minutes per school year in online math instruction (less than 30 minutes per week) compared to 1,500 minutes or more (42 minutes per week) for students at the 80th percentile. This trend can be partially explained by the provision of small-group tutoring within the learning lab to struggling math learners (approximately 10%–15% of students) identified by their teachers while other students worked on the online programs. These students had fewer opportunities during the year to work in an online instructional environment during regular lab time compared to the other students.[51] However, this trend may also be evidence of less engagement or productivity in the learning lab for low-performing math learners, although the research team collected no evidence during its site visits to validate this as a likely alternative explanation.

The finding that higher prior achievement was associated with more time spent in the online learning environments did not hold for all schools. In RLS, all third-grade students were active for similar amounts of time regardless of prior academic performance. In RDP, the relationship was in the opposite direction; students with the lowest prior academic performance scores spent more time active in online instructional programs compared to students with higher prior test scores. The differences across schools likely reflects the discretion schools had in how they organized students' time in the learning lab that may have resulted in different patterns in the use of online programs for different students.

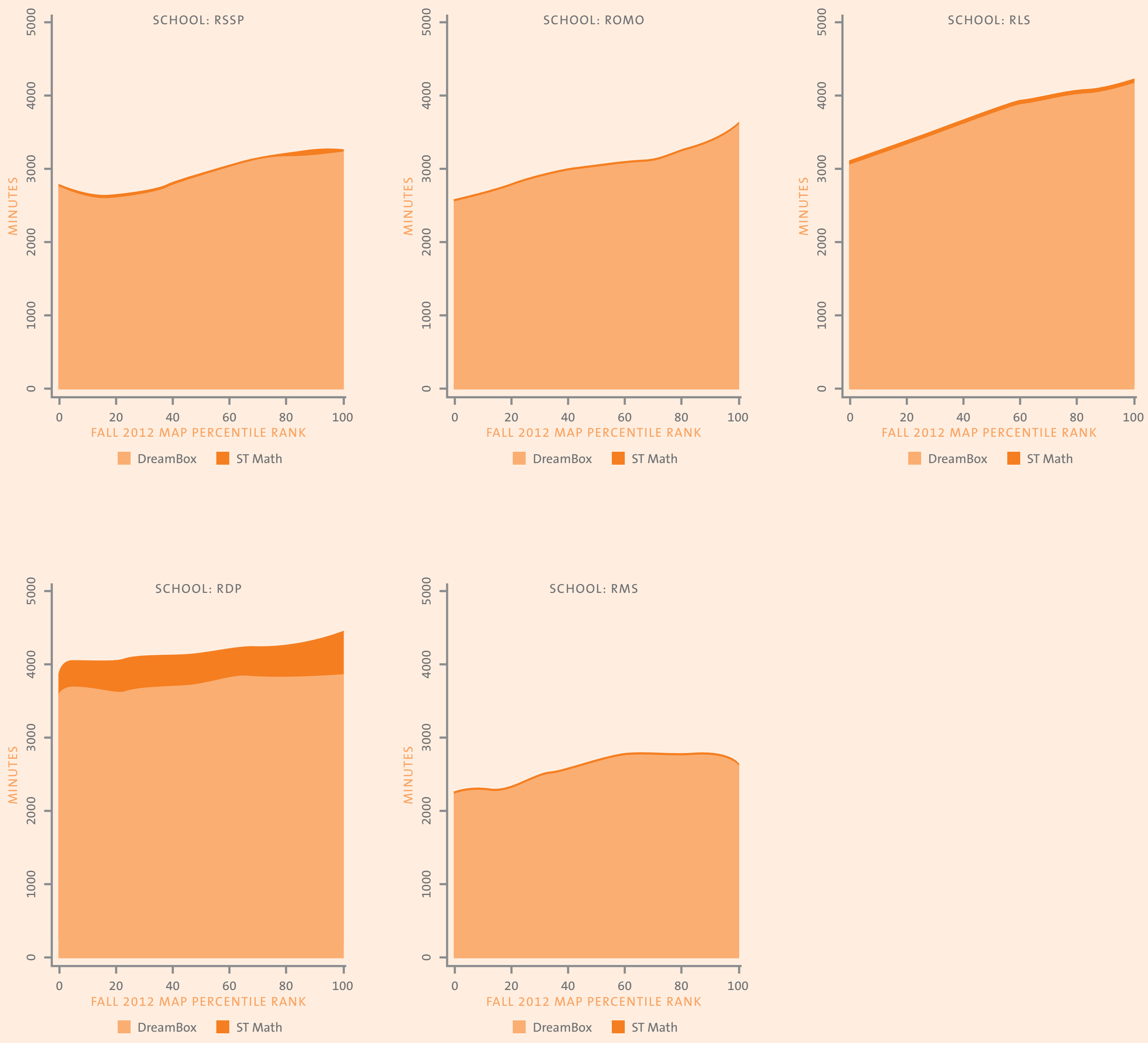
FIGURE 1. DISTRIBUTION OF TIME BY SCHOOL ACROSS GRADES



Note: (RDP = Discovery Prep, RLS = Los Suenos Academy, RMS = Mateo Sheedy Elementary ROMO = Mosaic Elementary, RSSP = Si Se Puede Academy)

[51] *Many of these students received their online math instruction in the morning before school (7:15 to 7:45 a.m.) or during the "enrichment" portion of the learning lab, when the other students were outside on the playground.*

FIGURE 2. DISTRIBUTION OF TIME BY PRIOR ACHIEVEMENT MEASURE BY SCHOOL, GRADE 1.



Note: Percentiles were generated separately for each school and grade level. The percentiles are not equivalent to national norms. (RDP = Discovery Prep, RLS = Los Suenos Academy, RMS = Mateo Sheedy Elementary ROMO = Mosaic Elementary, RSSP = Si Se Puede Academy) .

FIGURE 3. DISTRIBUTION OF TIME BY PRIOR ACHIEVEMENT MEASURE BY SCHOOL, GRADE 3.

Note: Percentiles were generated separately for each school and grade level. The percentiles are not equivalent to national norms. (RDP = Discovery Prep, RLS = Los Suenos Academy, RMS = Mateo Sheedy Elementary ROMO = Mosaic Elementary, RSSP = Si Se Puede Academy)

TABLE 8. OVERALL DIFFERENCE BETWEEN ROCKETSHIP AND THE VIRTUAL COMPARISON GROUP (VCG) ON SPRING NWEA MAP MATHEMATICS AND ELA SCORES

	% STUDENTS MEETING OR EXCEEDING SPRING SCORES FOR VCG STUDENTS	MEAN DIFFERENCE FROM VCG	STANDARD DEVIATION (SD) OF MAP SCORES	N	EFFECT SIZE	P
MAP MATH VCG						
Grade K	85.5	10.6	12.8	580	0.83	<0.001
Grade 1	70.4	5.7	14.0	463	0.41	<0.001
Grade 2	73.1	4.2	12.6	417	0.34	<0.001
Grade 3	68.3	3.4	13.4	407	0.26	<0.001
Grade 4	79.35	7.5	15.6	261	0.48	<0.001
Grade 5	88.9	12.2	15.1	90	0.81	<0.001
				Average	**0.52**	
MAP ELA VCG						
Grade K	68.6	4.8	13.2	574	0.36	<0.001
Grade 1	62.6	3.0	14.8	449	0.20	0.092
Grade 2	59.9	1.8	13.7	397	0.13	0.003
Grade 3	65.2	3.6	14.6	402	0.24	0.002
Grade 4	79.4	7.1	13.6	252	0.53	<0.001
Grade 5	76.7	8.2	12.6	90	0.65	<0.001
				Average	**0.35**	

Impact Analyses

IMPACT DESIGN OVERVIEW

A quasi-experimental design was used to explore the relationship between Rocketship's blended learning model use and gains in student learning. The design involved comparing the learning outcomes for students in the five Rocketship schools during SY 2011–12 to a comparison group of similar students attending similar schools that, ideally, did not adopt a blended learning model. Specifically, a "virtual comparison group" (VCG) was created for us by the Northwest Evaluation Association (NWEA) for comparison purposes using fall 2011 scores on NWEA's MAP assessment to identify similar students in similar schools who took the same test during the same test window (for additional details on the VCG analysis see the description in the main body of the report under Analytical Models section, page 17). Table A.1 in the Appendix for this profile shows the baseline descriptive statistics for students included in the sample for the impact analysis.

One of the primary limitations of an impact analysis using a "virtual comparison group" design is that the design is unable to isolate the effect of the blended learning model from other aspects of the schools' instructional system, including extended days and instructional time, as well as the academic culture of the school and school leadership. Blended learning is a critical component of Rocketship's instructional system and thus any differences in academic performance between Rocketship schools and other schools are at least partially explained by the adoption of blended learning.[52] However, it is not possible using a VCG design to estimate the relative importance of the blended learning model in comparison to other aspects of Rocketship's instructional system (including extended days and instructional blocks), teacher quality, school culture, or leadership.

[52] *Some schools in the virtual comparison group may also be using blended learning. Therefore, any effect that is directly related to Rocketship's blended learning model may underestimate the true effect of blended learning instruction relative to teacher-led instruction only.*

TABLE 9. DIFFERENCE BETWEEN ROCKETSHIP AND THE VIRTUAL COMPARISON GROUP ON SPRING NWEA MAP MATHEMATICS AND ELA SCORES, BY GENDER

	% STUDENTS MEETING OR EXCEEDING SPRING SCORES FOR VCG STUDENTS	MEAN DIFFERENCE FROM VCG	STANDARD DEVIATION (SD) OF MAP SCORES	N	EFFECT SIZE	P
MAP MATH VCG						
Grade K						
Female	85.4	10.2	12.8	280	0.79	<0.001
Male	85.7	11.0	12.8	300	0.86	<0.001
Grade 1						
Female	69.8	5.6	14.0	235	0.40	<0.001
Male	71.1	5.8	14.0	228	0.42	<0.001
Grade 2						
Female	76.4	4.6	12.6	216	0.37	<0.001
Male	69.7	3.8	12.6	201	0.30	<0.001
Grade 3						
Female	66.2	3.1	13.4	210	0.23	<0.001
Male	70.6	3.8	13.4	197	0.28	<0.001
Grade 4[A]						
Female	75.45	6.1	15.6	122	0.39	0.003
Male	82.7	8.7	15.6	139	0.56	<0.001
Grade 5						
Female	87.5	12.1	15.1	48	0.80	<0.001
Male	90.55	12.3	15.1	42	0.81	<0.001
MAP ELA VCG						
Grade K						
Female	69.7	4.8	13.2	274	0.36	<0.001
Male	67.7	4.7	13.2	300	0.36	<0.001
Grade 1						
Female	65.0	3.5	14.8	226	0.24	0.051
Male	60.1	2.4	14.8	223	0.16	0.186
Grade 2						
Female	58.4	1.7	13.7	209	0.12	0.030
Male	61.7	2.0	13.7	188	0.15	0.011
Grade 3						
Female	68.1	3.7	14.6	207	0.25	0.003
Male	62.1	3.4	14.6	195	0.23	0.007
Grade 4[A]						
Female	73.3	5.2	13.6	120	0.38	<0.001
Male	84.8	8.8	13.6	132	0.65	<0.001
Grade 5						
Female	75.0	7.8	12.6	48	0.62	<0.001
Male	78.6	8.7	12.6	42	0.69	<0.001

[A] *Exhibited statistically significant gender differences in the MAP VCG effect sizes ($p < .05$).*

Michael & Susan Dell FOUNDATION

TABLE 10. DIFFERENCE BETWEEN ROCKETSHIP AND THE VIRTUAL COMPARISON GROUP ON SPRING NWEA MAP MATHEMATICS AND ELA SCORES, BY PRIOR ACHIEVEMENT

	% STUDENTS MEETING OR EXCEEDING SPRING SCORES FOR VCG STUDENTS	MEAN DIFFERENCE FROM VCG	STANDARD DEVIATION (SD) OF MAP SCORES	N	EFFECT SIZE	P
MAP MATH VCG						
Grade K[A]						
Higher Prior	93.1	12.0	12.8	291	0.94	<0.001
Lower Prior	77.9	9.3	12.8	289	0.72	<0.001
Grade 1[A]						
Higher Prior	67.1	4.6	14.0	228	0.33	0.001
Lower Prior	73.6	6.7	14.0	235	0.48	<0.001
Grade 2						
Higher Prior	78.4	4.5	12.6	199	0.35	<0.001
Lower Prior	68.3	4.0	12.6	218	0.32	<0.001
Grade 3						
Higher Prior	72.1	3.3	13.4	208	0.25	<0.001
Lower Prior	64.3	3.6	13.4	199	0.27	<0.001
Grade 4[A]						
Higher Prior	83.9	6.9	15.6	124	0.44	0.001
Lower Prior	75.2	8.1	15.6	137	0.52	<0.001
Grade 5						
Higher Prior	83.7	9.0	15.1	43	0.59	<0.001
Lower Prior	93.6	15.1	15.1	47	1.00	<0.001
MAP ELA VCG						
Grade K[A]						
Higher Prior	76.1	6.5	13.2	293	0.49	<0.001
Lower Prior	60.9	3.0	13.2	281	0.23	0.001
Grade 1						
Higher Prior	63.4	2.5	14.8	227	0.17	0.173
Lower Prior	61.7	3.5	14.8	222	0.24	0.057
Grade 2						
Higher Prior	61.8	1.7	13.7	191	0.13	0.028
Lower Prior	58.3	1.9	13.7	206	0.14	0.013
Grade 3						
Higher Prior	68.8	3.6	14.6	202	0.25	0.004
Lower Prior	61.5	3.5	14.6	200	0.24	0.005
Grade 4[A]						
Higher Prior	75.0	4.3	13.6	132	0.32	0.001
Lower Prior	84.2	10.2	13.6	120	0.75	<0.001
Grade 5[A]						
Higher Prior	65.9	3.5	12.6	44	0.27	0.015
Lower Prior	87.0	12.8	12.6	46	1.02	<0.001

[A] *Exhibited statistically significant prior achievement differences in the MAP VCG effect sizes ($p < .05$).*

IMPACT FINDINGS

Rocketship students outperformed students in the VCG in math and ELA across all grade levels (Table 8). In math, between 68% (Grade 3) and 89% (Grade 5) of Rocketship students met or exceeded the gains made by the students in the virtual comparison group. In English language arts (ELA), between 60% (Grade 2) and 79% (Grade 4) of the students met or exceeded the gains made by the students in the VCG. The differences between Rocketship schools and the VCG across all grade levels are statistically significant. On average, the effects are 50% larger for math (average effect size = +0.52) than for reading (average effect size = +0.35).

There is also evidence that performance of students in Grade 4 relative to the VCG varied by gender in both math and ELA (see Table 9). Specifically, fourth-grade male students outgained female students relative to the VCG in math (83% vs. 75%) and ELA (85% vs. 73%). There were no other statistically significant differences by gender across the other grade levels.

We also found evidence that performance relative to VCG may vary by prior achievement for both math and ELA across several grade levels (see Table 10). For math, kindergarten students with higher prior achievement levels (at or above the median score on the fall MAP) outgained their peers with lower prior achievement relative to the VCG (93% vs. 78%). For first and fifth-grade students, however, students with lower prior achievement levels significantly outperformed students with higher prior achievement levels relative to the VCG (74% vs. 67% and 93% vs. 83%, respectively). For ELA, we see a similar pattern of results, with kindergarten students with higher prior achievement levels significantly outgaining students with lower prior achievement levels relative to the VCG (76% vs. 61%), but in fifth grade, students with lower prior achievement levels significantly outperformed students with higher prior achievement levels (87% vs. 66%). In addition, in fourth grade, students with lower prior achievement in reading outgained students with higher prior achievement on the spring reading MAP test relative to the VCG (84% vs. 75%).

ACHIEVEMENT GAP ANALYSIS

In addition to using the virtual comparison design to estimate the relative effect of the Rocketship instructional model, we also examined the extent to which Rocketship's schools were narrowing the achievement gap between high-income and low-income schools on the California state standardized achievement tests (California Standards Test or CST). The results of achievement gap analysis for Rocketship schools are presented below for spring 2012 second- and third-grade ELA and math test scores (the two grade levels for which state test scores are available and which present in all schools). Table 11 shows the sample size for the different groups included in the achievement gap analysis for Rocketship. We define high-income schools as public high schools in California with 10% or fewer of its students qualifying for the federal free or reduced-price lunch program. Low-income schools are those schools in the state whose proportion of students who qualify for free or reduced-price lunch are within plus or minus 5% of the Rocketship schools. Approximately 90% of Rocketship's students qualified for the federal free or reduced-priced lunch program.

TABLE 11: SAMPLE SIZE FOR THE ACHIEVEMENT GAP ANALYSIS

	ELA	MATH
Rocketship schools	5	5
Low-income comparison schools	1,040	1,040
High-income comparison schools	475	475

We used two different approaches to estimate the reduction in the achievement gap in ELA and math between high-income and Rocketship schools in state achievement scores—a comparison of performance levels and scale point scores (for details on the achievement gap analyses see description above in the main body of the report in the Analytical Models section, page 17). Results of our analysis showed that the Rocketship schools were significantly reducing the high- versus low-income achievement gap in Grades 2 and 3. The reduction was greater for math scores than for ELA (approximately 75% of Rocketship students were designated as English language learners) and greater for Grade 2 than Grade 3. The results are summarized in Table 12 (ELA) and Table 13 (math). Figures 4–7 further illustrate the degree of the achievement gap reduction across subject areas and grade levels.

Comparison of performance levels. As shown in Figure 4 (Grade 2) and Figure 5 (Grade 3), the average difference, or gap, between high- and low-income schools in the percentage of students in Grades 2 and 3 in California scoring at or above proficient based on spring 2012 CST scores between high-income and low-income schools in California varied from 24% in math (Grade 3) to 39% in ELA (Grade 3). When we examined the achievement gap between high-income elementary schools in California and Rocketship schools we found that the gap was

TABLE 12: SUMMARY OF THE ACHIEVEMENT GAP ANALYSIS RESULTS FOR ENGLISH LANGUAGE ARTS (ELA)

ELA, CALIFORNIA STANDARDS TEST(CST)		ACHIEVEMENT GAP		
	GRADE LEVEL	HIGH-INCOME VS. LOW-INCOME SCHOOLS	HIGH-INCOME VS. ROCKETSHIP SCHOOLS	% REDUCTION IN ACHIEVEMENT GAP
% STUDENTS SCORING AT OR ABOVE PROFICIENT				
	2	33%	14%	58%
	3	39%	27%	32%
SCALE SCORES				
	2	56	27	52%
	3	59	43	27%

TABLE 13: SUMMARY OF THE ACHIEVEMENT GAP ANALYSIS RESULTS FOR MATH

MATH, CALIFORNIA STANDARDS TEST (CST)		ACHIEVEMENT GAP		
	GRADE LEVEL	HIGH-INCOME VS. LOW-INCOME SCHOOLS	HIGH-INCOME VS. ROCKETSHIP SCHOOLS	% REDUCTION IN ACHIEVEMENT GAP
% STUDENTS SCORING AT OR ABOVE PROFICIENT				
	2	29%	1%	97%
	3	24%	10%	58%
SCALE SCORES				
	2	71	-5	107%
	3	65	37	43%

reduced and ranged from 1% for math (Grade 2) and 27% in ELA (Grade 3). This represented a reduction in the income-based achievement gap of 97% in Grade 2 math and 32% for Grade 3 ELA.

Comparison of scale scores. When the comparison is CST scale point scores, the estimated reduction in the high-income/low-income achievement gap is similar but even greater for Grade 2 math. For ELA CST scores, enrollment in a Rocketship school was associated with a 52% reduction in the achievement gap in Grade 2 and a 27% reduction in Grade 3. For math, Rocketship enrollment is associated with 107% reduction in Grade 2 (on average Rocketship students outscored their high-income peers) and 58% in math.

While the reduction in the high-income/low-income achievement gap associated with attending a Rocketship school is significant, caution must be taken in interpreting these results as a function of Rocketship's blended learning model alone. This analysis uses a quasi-experimental comparison design that does not control for preexisting differences that might exist between families and students who enroll in Rocketship schools and those who enroll in the average low-income school in the state. While these differences may contribute to some of the reduction in the achievement gap, this contribution is likely to be small. We have no direct evidence that Rocketship schools are recruiting a higher than expected proportion of more advantaged families from the communities around their schools (90% of Rocketship students qualified for the federal subsidized lunch program and 75% of students were designated as English language learners), although only 4% of Rocketship students have an Individualized Education Plan (IEP), a formal indication of a special learning need, compared to about 12% in the local public school district (San Jose Unified School District).

FIGURE 4: DIFFERENCE IN PERCENT AT OR ABOVE PROFICIENT ON CALIFORNIA STANDARDS TEST (CST) BETWEEN ROCKETSHIP SCHOOLS AND HIGH-INCOME AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA, GRADE 2.

FIGURE 5: DIFFERENCE IN PERCENT AT OR ABOVE PROFICIENT ON CALIFORNIA STANDARDS TEST (CST) BETWEEN ROCKETSHIP SCHOOLS AND HIGH-INCOME AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA, GRADE 3.

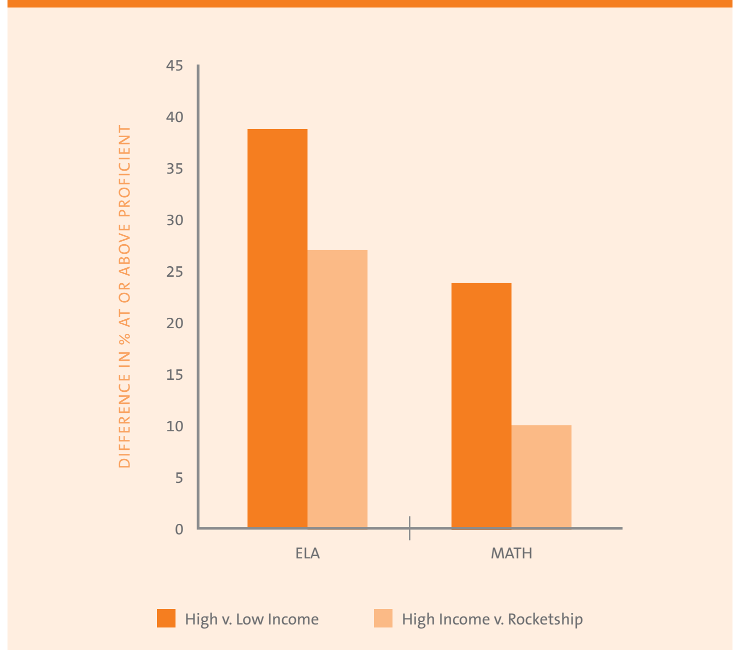
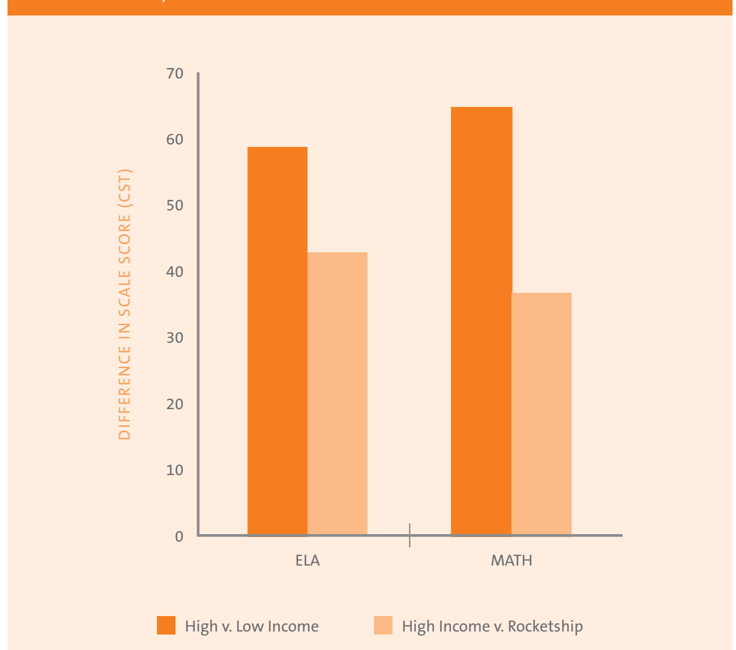


FIGURE 6: DIFFERENCE IN CALIFORNIA STANDARDS TEST (CST) SCALE SCORE POINTS BETWEEN ROCKETSHIP SCHOOLS AND HIGH-INCOME AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA, GRADE 2.

FIGURE 7: DIFFERENCE IN CALIFORNIA STANDARDS TEST (CST) SCALE SCORE POINTS BETWEEN ROCKETSHIP SCHOOLS AND HIGH-INCOME AND LOW-INCOME COMPARISON SCHOOLS IN CALIFORNIA, GRADE 3.



Michael & Susan Dell FOUNDATION

Appendix: Rocketship Education

TABLE A.1. BASELINE DESCRIPTIVE STATISTICS: STUDENTS IN THE ANALYTIC SAMPLE, GRADES K–5

VARIABLE	PERCENT	N
ETHNICITY		
American Indian / Alaska Native	1	12
Asian	10	234
African American	2	52
Hispanic	85	2,132
Other/Declined	.5	6
Pacific Islander	1	10
White	3	75
GENDER		
Female	50	1,251
Male	51	1,275
ENGLISH LANGUAGE LEARNER (ELL) STATUS		
ELL	72	1,643
Not ELL	33	680
ELIGIBLE FOR FEDERAL LUNCH PROGRAM		
Yes	84	2,123
No	16	397
INDIVIDUALIZED EDUCATION PLAN (IEP) (SPECIAL EDUCATION STATUS)		
IEP	4	86
No IEP	96	2,331

Teacher Survey Results – Rocketship Education

SURVEY ADMINISTRATION

Overall, 51 of 70 Rocketship teachers implementing blended learning completed the Teacher Survey for a 73% response rate. Table S1 describes the characteristics of the respondents in terms of the subject area(s) they teach and years of teaching experience. Additionally, 24 of 28 lab monitors supporting the learning labs completed a Lab Monitor Survey for an 85% response rate.

TABLE S1. TEACHER SURVEY RESPONDENT CHARACTERISTICS

TEACHER CHARACTERISTICS	NUMBER	PERCENT
SUBJECT TAUGHT		
English Language Arts (ELA)	34	67
Math	17	33
Math and ELA	1	2
TEACHING EXPERIENCE		
New teachers (3 years or less)	29	57
Veteran teachers (4 or more years)	22	43

FIGURE S1. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES (ELA CLASSROOMS).

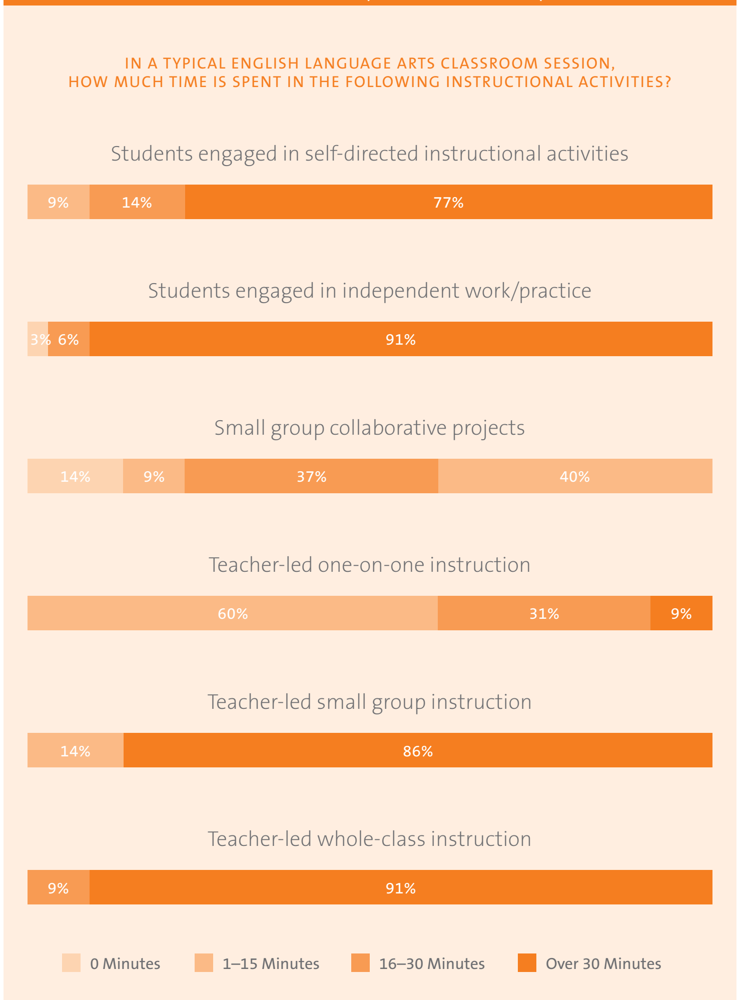
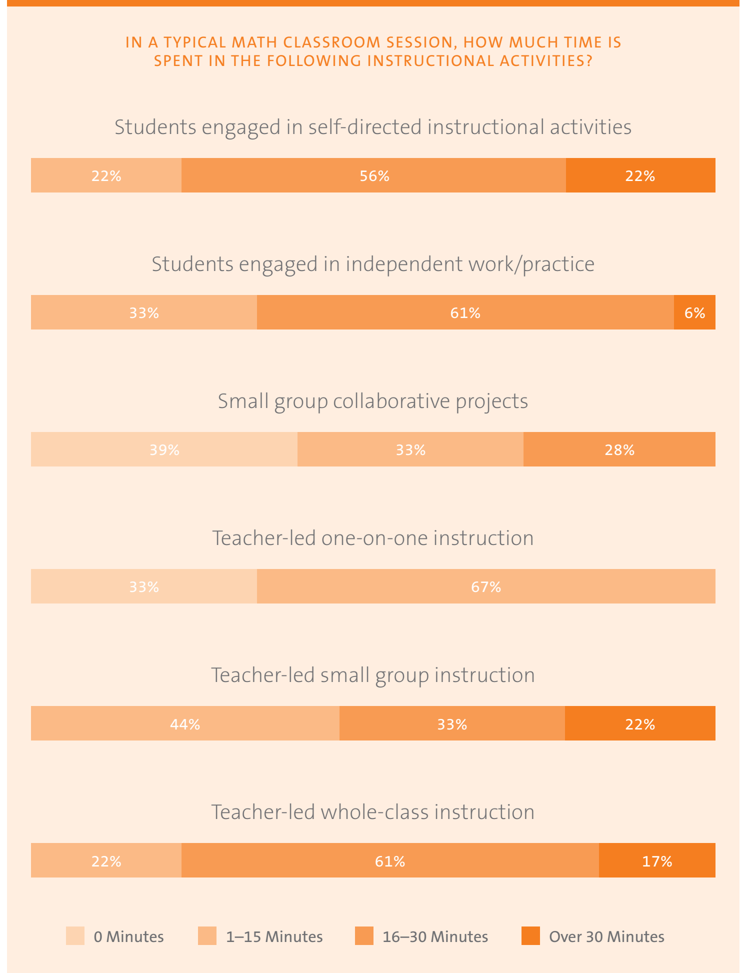


FIGURE S2. PERCENTAGE OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS (MATH CLASSROOMS).



Classroom Instructional Activities

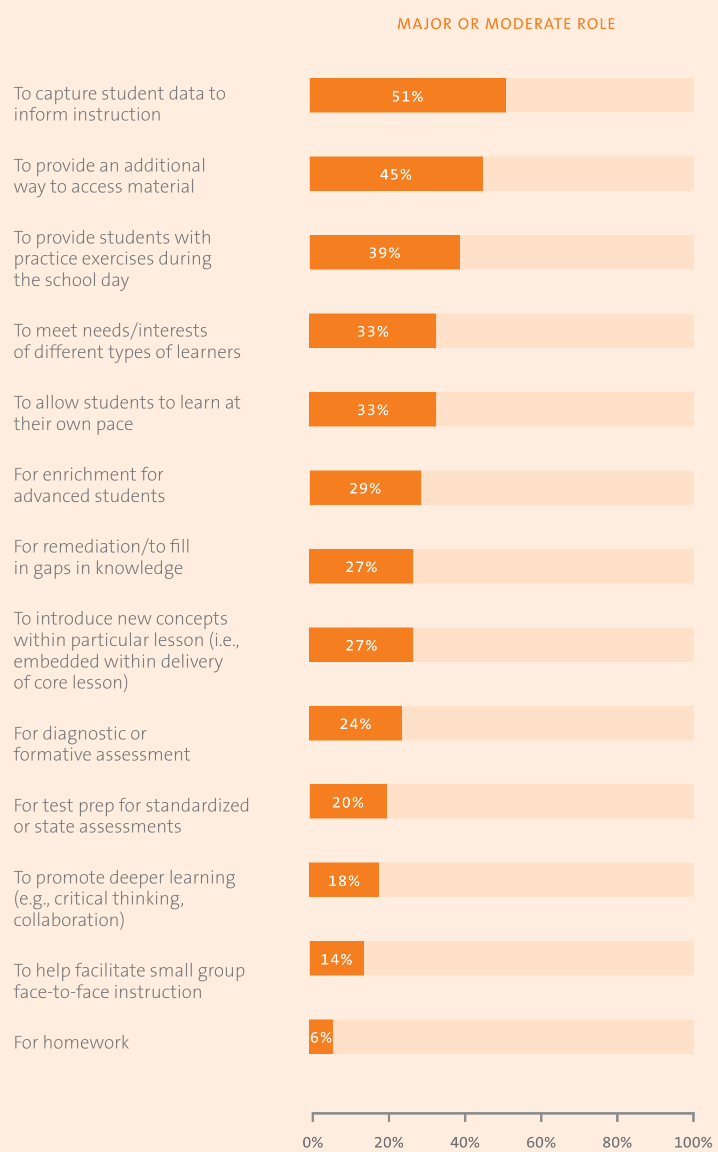
ENGLISH LANGUAGE ARTS

Rocketship's English language arts (ELA) teachers reported devoting over 30 minutes of a typical ELA classroom session to teacher-led whole-class instruction (91%), independent work/practice (91%), teacher-led small-group instruction (86%), and student-directed instructional activities (77%), with smaller proportions of teachers devoting the same amount of time to small-group collaborative projects (40%) and teacher-led one-on-one instruction (9%) (see Figure S1).

MATH

Rocketship math teachers reported devoting between 16–30 minutes of a typical math classroom session to teacher-led whole-class instruction and independent work/practice (61% for both), and student-directed instructional activities (56%), with smaller proportions of teachers devoting the same amount of time to teacher-led small-group instruction (33%) and small-group collaborative projects (28%). These math teachers also report devoting over 30 minutes to teacher-led small-group instruction and student-directed instructional activities (22% for both) (see Figure S2).

FIGURE S3: PERCENT OF TEACHERS REPORTING DIFFERENT ROLES FOR TECHNOLOGY AND WEB-BASED INSTRUCTION IN SUPPORTING TEACHER-PROVIDED INSTRUCTION.



ROLE OF TECHNOLOGY

Rocketship teachers reported that technology and web-based instruction played a major or moderate role in supporting their instruction by helping them capture student data to inform instruction (51%), and by providing students with an additional way to access material (45%), and providing practice exercises during the school day (39%) (see Figure S3).

Generally speaking, technology and web-based instruction played a greater role in math than in ELA in a number of areas, such as providing practice exercises for remediation/enrichment, for test prep, for diagnostic assessment, and to personalize learning with half or more of the math teachers indicating it played a major or moderate role in these areas compared to about a fourth or less of ELA teachers. In contrast, technology was used to a greater extent in ELA to introduce new concepts (32% vs. 13% in math).

THE LEARNING LAB

Lab monitors reported that their primary roles in supporting the learning lab revolved around lab management (e.g., keeping students on task), motivation and engagement (e.g., acknowledging good behavior), and student learning (e.g., tutoring), and that secondary roles included managing students' use of technology (e.g., helping with log in information) and setting student goals with teachers.

Only half (52%) of lab monitors reported that they had access to real-time reports on student progress and/or desktop monitoring software that allowed them to monitor whether or not students were on task while they were in the lab; of those, 31% used it every day or almost every day, 8% used it two or three times a week, 54% used it about once a week, and 8% never used it.

Rocketship teachers had minimal contact with the students in the learning lab—65% reported that they never supervised students in the learning lab. Likewise, Rocketship teachers had minimal contact with the learning lab monitors, with close to half (46%) meeting with monitors once a month or less (see Table S2). ELA teachers reported meeting with lab monitors on a more frequent basis than math teachers to discuss student performance in the learning lab, with a quarter (24%) of math

TABLE S2. TEACHER REPORTS OF FREQUENCY IN WHICH THEY MEET WITH LAB MONITORS

HOW OFTEN DO YOU DO MEET WITH LAB MONITORS TO DISCUSS STUDENT PERFORMANCE IN THE LEARNING LAB?	ELA (%)	MATH (%)	OVERALL (%)
Never	0	24	**8**
About once a month or less	38	41	**38**
Two or three times a month	21	29	**25**
About once a week or more	42	6	**29**

teachers indicating they never did this. Instead, 84% of lab monitors reported meeting with the assistant principal about once a week or more to discuss student performance in the learning lab.

The most frequently cited topic of discussion between teachers and lab monitors was students' behavior in the learning lab, followed by students' progress on tutoring materials, students' understanding of specific concepts, and students' progress or performance goals. Both teacher and lab monitors reported wanting to meet more frequently and that it was important that they know what was going on, particularly what students struggle with, in both the classroom (for lab monitors) and the learning lab (for teachers) to be effective in their respective roles.

Rocketship teachers were split on the extent to which there was alignment between the classroom and learning lab curricula—just over one-third of teachers reported that the learning lab material was moderately aligned or very well aligned with the material they presented in the classroom, while another one-third of teachers reported that the learning lab material was only weakly aligned with the material they presented in the classroom. The remaining one-third reported that they don't know how well the materials were aligned, most likely due to the limited time they spent in the learning lab (see Table S3).

TABLE S3. ALIGNMENT OF CLASSROOM AND LEARNING LAB CURRICULUM

IN YOUR OPINION, HOW WELL ALIGNED IS THE MATERIAL STUDENTS ARE EXPOSED TO IN THE LEARNING LAB WITH THE MATERIAL YOU PRESENT IN THE CLASSROOM?	ELA (%)	MATH (%)	OVERALL (%)
Very well aligned	6	18	**10**
Moderately aligned	24	41	**29**
Weakly aligned	41	12	**33**
Not at all aligned	3	0	**2**
Don't know	26	29	**27**

Math teachers reported stronger alignment between the online materials students were exposed to in the learning lab and the material they presented in the classroom; more than half (59%) felt it was moderately or very well aligned, compared with only a third (30%) of ELA teachers.

USE OF DATA

Approximately three-quarters (71%) of teachers reported that they reviewed student performance data recorded by the software system(s). Of those who reviewed the data, more than one-half reviewed the data at least once a week, while one in three teachers looked at the data about once a month or less (see Table S4). Math teachers reported looking at the student performance data captured by the program(s) on a less frequent basis than ELA teachers, with many (73%) reviewing it once a month or less.

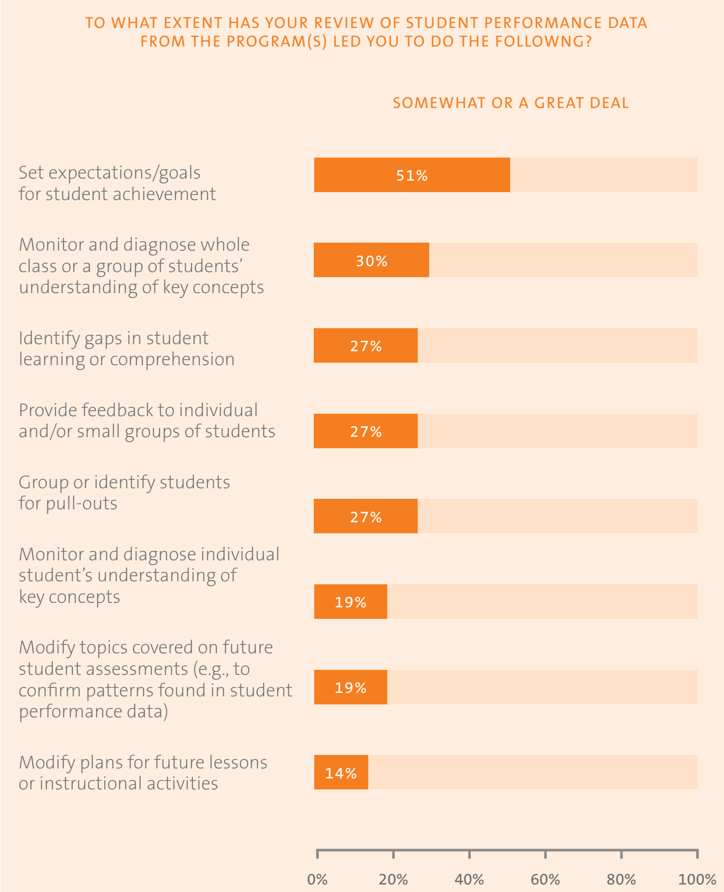
TABLE S4. FREQUENCY OF DATA REVIEW BY **ROCKETSHIP** TEACHERS

HOW OFTEN DO YOU LOOK AT THE DATA RECORDED BY THE SYSTEM(S) USED TO SUPPORT BLENDED LEARNING?	ELA (%)	MATH (%)	OVERALL (%)
Every day or almost every day	44	0	**31**
Two to three times a week	12	0	**8**
About once a week	16	18	**17**
Two to three times a month	16	9	**14**
About once a month or less	12	73	**31**

Teachers cited multiple reasons for not reviewing the data at all, or once a month or less, including not having a way to easily access reports of student performance data, relying more on information outside the software systems, and not understanding how to use the data to inform their instruction.

Almost two-thirds (60%) of teachers reported that the data was only slightly useful or not at all useful for informing their instruction. Despite the differences in their frequency of review, math and ELA teachers provided similar reports regarding the usefulness of the data for informing their instruction. Teachers that reviewed the data reported using it to set expectations/goals for student achievement; monitor, diagnose, and provide feedback on whole-class and group understanding of key concepts; and identify gaps in student learning or comprehension (see Figure S4).

FIGURE S4. PERCENT OF TEACHERS USING DATA "SOMEWHAT" OR "A GREAT DEAL" FOR DIFFERENT INSTRUCTIONAL PURPOSES.



ELA teachers indicated greater use of the data from software program(s) to set expectations and goals for student achievement than math teachers; 65% of ELA teachers indicated the data led them to do this somewhat or a great deal, compared to only 18% of math teachers. In many of the other areas listed, however, math teachers reported more use of the data, including to modify plans for future lessons or assessments, grouping students for pull-outs, and monitoring the understanding of the class as a whole on key concepts; between a third and a half of math teachers reported the data led them to do these things somewhat or a great deal, compared to a quarter or less of the ELA teachers.

Additionally, 88% of lab monitors indicated they reviewed the student performance data recorded by the software system(s) used to support blended learning, typically on at least a weekly basis if not more. The majority (91%) of lab monitors reported that the student performance data was somewhat useful or very useful for supporting their monitoring of the learning lab. Rocketship lab monitors reported that reviewing student performance data from program(s) led them to identify students who were not making sufficient progress, to set expectations/goals for student achievement, and to identify gaps in student learning or comprehension.

Satisfaction with Blended Learning and Impact on Student Learning

TEACHER SATISFACTION WITH BLENDED LEARNING

Overall, teachers appeared to be only moderately satisfied with the role of blended learning in instruction. Only one-half of teachers would recommend blended learning to other teachers (see Table S5). Reported levels of satisfaction were reasonably consistent across math and ELA teachers; however, more of the ELA teachers agreed that the use of the blended model helped improve students' learning and understanding (51% vs. 31% of math teachers).

TABLE S5. PERCENTAGE OF TEACHERS REPORTING THEY "AGREE" OR "STRONGLY AGREE" THAT BLENDED LEARNING BENEFITS THEIR STUDENTS IN DIFFERENT WAYS.

PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS.	PERCENT AGREE OR STRONGLY AGREE		
	ELA (%)	MATH (%)	OVERALL (%)
The learning lab model meets the learning needs of my students.	47	56	49
Students are highly engaged while using the learning lab model.	56	44	51
The learning lab model helps students take ownership for their own learning.	35	31	33
Students' learning and understanding of the materials has improved due to the use of the learning lab model.	51	31	44
I would recommend the use of the learning lab model to other teachers.	56	44	51

FIGURE S5: PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF EFFECTIVENESS OF THE BLENDED LEARNING MODEL FOR STUDENTS WITH DIFFERENT LEVELS OF ACADEMIC PERFORMANCE OR NEEDS.

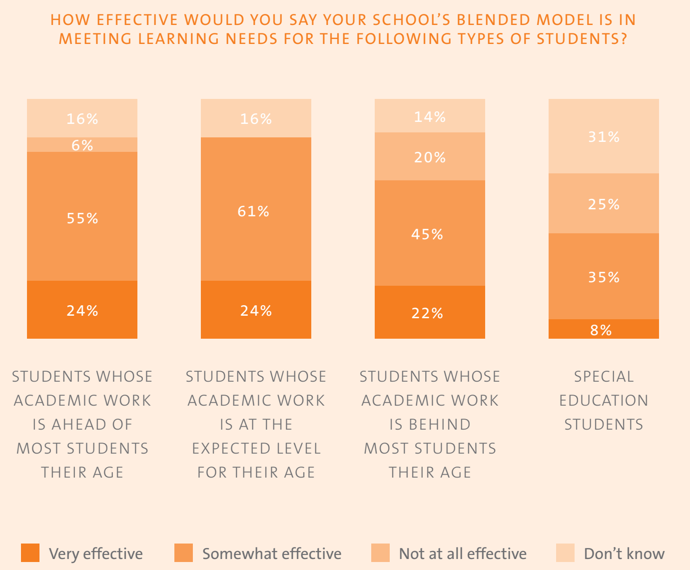
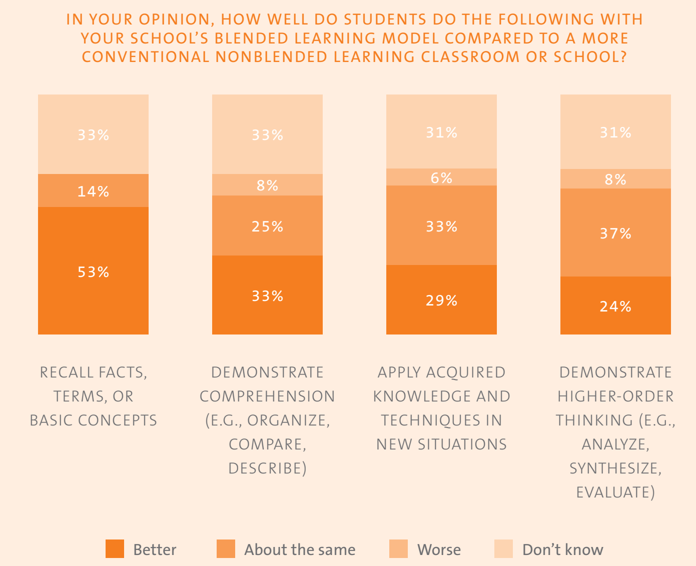


FIGURE S6. PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF THE EXTENT TO WHICH THEIR STUDENTS DEMONSTRATE HIGH-ORDER THINKING AS A RESULT OF THE BLENDED LEARNING MODEL RELATIVE TO A MORE CONVENTIONAL INSTRUCTIONAL MODEL.



EFFECTIVENESS FOR STUDENT TYPES

According to teacher reports, the learning lab model may be most effective for students whose academic work was ahead of most students their age, or at the expected level for their age, and least effective for students whose academic work was behind most students their age and for special education students. Specifically, 24% of Rocketship teachers report that their blended learning model was very effective for students whose academic work was ahead of, or at the expected level for, most students their age. This dropped to 22% for students whose academic work was behind most students their age and to 8% for special education students (see Figure S5). Teachers' ratings of the effectiveness of blended learning for various types of students were consistent across subject area.

EFFECTIVENESS RELATIVE TO NONBLENDED ENVIRONMENT

Teachers reported that blended learning had likely improved students' ability to recall basic facts more than any other skill relative to nonblended learning environments (see Figure S6). Specifically, Rocketship teachers reported that students participating in blended learning recall facts, terms, or basic concepts (53%), demonstrate comprehension (33%), apply acquired knowledge and techniques in new settings (29%), and demonstrate higher-order thinking (24%) better than students in more traditional nonblended classrooms or schools.

Regardless of their years of teaching experience, teachers at Rocketship provided reasonably similar ratings when asked to compare students' skills in various areas under their school's blended learning model versus a more traditional, nonblended environment; roughly the same percentage of new and veteran teachers indicated students did the same or better with blended learning. However, a handful of new teachers at Rocketship indicated students did worse with blended learning, while more of the veteran teachers indicated they did not know. Teachers' ratings of students' skills in a variety of areas with blended learning compared to a traditional nonblended environment were consistent across subject area.

TABLE S6. PERCENTAGE OF TEACHERS REPORTING DIFFERENT LEVELS OF SATISFACTION WITH TRAINING.

OVERALL, HOW SATISFIED ARE YOU WITH THE TRAINING YOU HAVE RECEIVED TO USE THE LEARNING LAB MODEL?	TEACHERS (%)	LAB MONITORS (%)
Very satisfied	0	**18**
Satisfied	39	**50**
Dissatisfied	55	**32**
Very dissatisfied	7	**0**

TRAINING AND SUPPORT

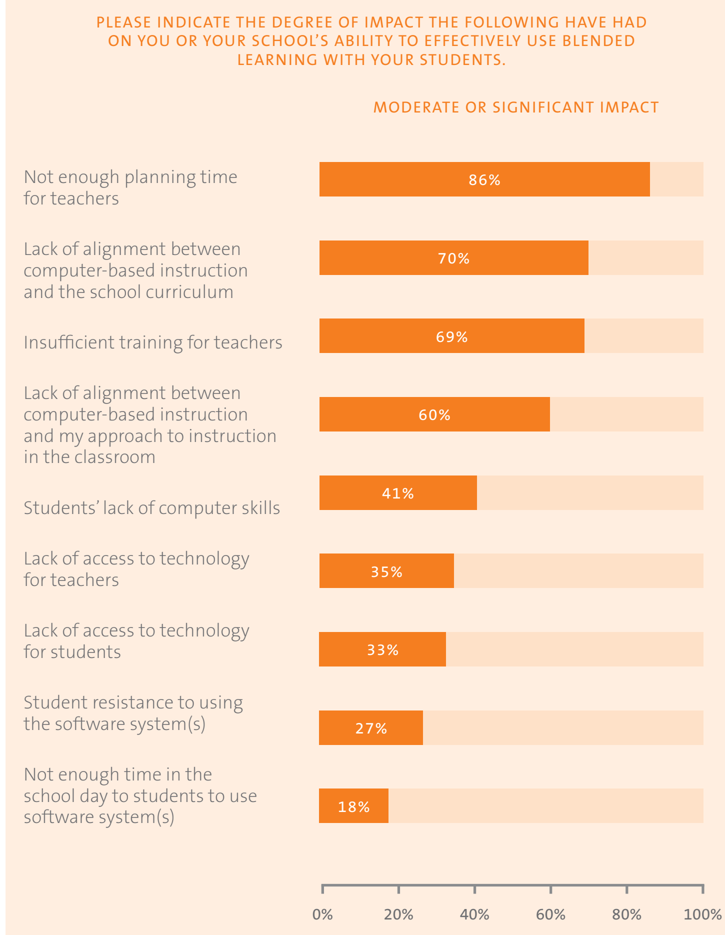
Most (86%) Rocketship teachers participated in a training or orientation session (either in person or online) directly related to their school's use of blended learning or on the specific software program(s) supporting it; those who did not either indicated no training was provided or they were not aware of any training being offered. Additionally, 84% of lab monitors participated in a training or orientation session.

Lab monitors were generally more satisfied with the training they received than teachers (see Table S6). A number of the Rocketship teachers who indicated they were not satisfied with the training said that they lacked information about what their students were doing in the lab (e.g., what content they were exposed to, how much time they were spending on different programs). Teachers also cited either a lack of data and/or access to it, as well as not enough information about how to use it to measure student progress and inform their instruction.

Half (53%) of teachers reported spending less than 1 hour of their own time (i.e., outside of professional development) getting acquainted with the software program(s) supporting the learning lab model, or planning for how to best integrate the learning lab model with their instruction; another 43% of teachers reported spending between 1–5 hours of their own time, and only 4% of teachers reported spending more than 5 hours of their own time devoted to these tasks. In contrast, only 16% of lab monitors reported spending less than 1 hour of their own time (i.e., outside of professional development) getting acquainted with the software program(s) supporting the learning lab model; 64% spent between 1–5 hours and 20% spent more than 5 hours.

Close to two-thirds (60%) of lab monitors reported that both technical/IT support and support from other staff (e.g., other facilitators, teachers, local school leadership, off-site administrative staff) was fully adequate for them to be an effective lab facilitator. Most of the remaining lab monitors reported the support in these areas as somewhat adequate. These reports were associated with IT support staff not being based on-site, inconsistent/lengthy response time to address problems, and not able to address questions about program content. Reports of only somewhat adequate support from other staff related to the need for stronger relationship with teachers and more exchange of information between the classroom and lab, and desire for increased presence of other staff in the lab so these individuals would have more firsthand knowledge about what goes on.

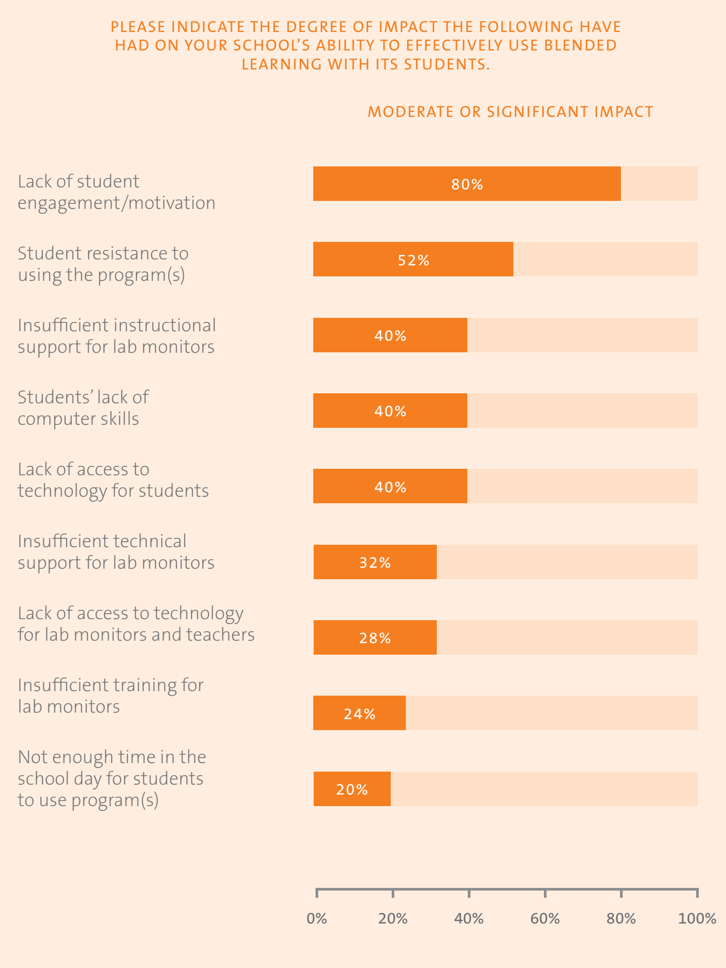
FIGURE S7: PERCENTAGE OF TEACHERS REPORTING THE EXTENT TO WHICH DIFFERENT FACTORS INFLUENCED THE EFFECTIVE USE OF THE BLENDED LEARNING MODEL TO A "MODERATE" OR "SIGNIFICANT" DEGREE.



FACTORS INFLUENCING USE

Rocketship teachers and lab monitors reported different sets of barriers regarding their school's ability to effectively use the learning lab model with their students (Figures S7 and S8). Teachers reported that not enough planning time, lack of alignment between learning lab instruction and school curriculum, insufficient training, and the lack of alignment between online and personal approaches to instruction all had moderate or significant impacts on their ability to effectively use blended learning with their students. On the other hand, a majority of lab monitors reported that lack of student

FIGURE S8: PERCENTAGE OF TEACHERS REPORTING THE EXTENT TO WHICH DIFFERENT FACTORS INFLUENCED THE EFFECTIVE USE OF THE BLENDED LEARNING MODEL TO A "MODERATE" OR "SIGNIFICANT" DEGREE.



engagement or motivation and student resistance to using programs had moderate or significant impacts on their school's ability to effectively use blended learning with their students.

More than two-thirds of Rocketship teachers reported experiencing, or their students experiencing, technical problems while using the learning lab model, and of these teachers, 6 in 10 teachers reported the technical problems being a moderate or significant barrier to effective use of the model (approximately 30% of teachers reported they did not know whether or not technical issues existed).

Nearly all (92%) lab monitors also reported experiencing, or their students experiencing, technical problems while using the learning lab model, and of these lab monitors, approximately half reported the technical problems being a moderate or significant barrier. The most commonly cited technical problems for both teachers and lab monitors included malfunctioning computers and unreliable Internet connectivity and software. In addition, over three-quarters of the lab monitors who reported the presence of technical issues also reported insufficient Internet bandwidth as a technical challenge they faced in the lab.

Michael & Susan Dell
FOUNDATION

Research Profile: Summit Public Schools (2011–12 School Year)

Introduction

Summit Public Schools (Summit), founded in 2003, is a charter management organization (CMO) that currently operates six high schools located in the San Francisco Bay Area. Summit's schools serve a diverse student population and recruit students from the low-income neighborhoods that surround the schools, as well as middle and upper middle class students looking for an alternative to their local public high school. Summit's mission is to "help every student enrolled in one of its schools to have an opportunity to attend, and succeed in, a four-year college or university, and to be thoughtful, contributing members of society."

In the 2011–12 school year, Summit opened two new schools that were co-located on the campus of the National Hispanic University, in San Jose, California. These schools adopted a blended learning model for their ninth-grade math instruction using Khan Academy to supplement the core classroom curriculum. Teachers and students used Khan Academy videos, practice exercises, and performance tracking and progress monitoring tools to supplement their normal instruction.

The decision to open the schools with blended learning math programs was motivated by two factors. First, Summit was hoping that Khan Academy could help incoming students with low math achievement and fill gaps in their basic skills that were a legacy of their primary school experience. Second, Summit leadership was seeking ways to create a more personalized learning experience for students to drive learning gains by allowing students to work on Khan Academy problem sets at their own pace while helping teachers facilitate small-group teacher-directed instruction and individual one-on-one tutoring.

BLENDED MODEL DESCRIPTION

During most of the 2011–12 school year, Summit Public School's Tahoma and Rainier schools in San Jose, California, used Khan Academy resources within a rotation blended learning model.[53] During the first 8 weeks of the school year, students used Khan Academy in the classroom but during a time that was segregated from core teacher-led math instruction. Each classroom had its own set of laptop computers and high-speed wireless Internet to support one-to-one computing. Students typically used the Khan Academy resources in the second half of their daily 90-minute math instruction block. Typical use was 30–60 minutes a day, Monday through Thursday (Friday is a half-day for students). During the first 4 weeks of the school year, students were also assigned to use Khan Academy at home for 60 minutes per week to become more familiar with the resources and features of the website. The duration of daily use varied by teacher and often depended on whether the teacher decided to use some of the time set aside for Khan Academy to extend and complete the lesson planned for the regular instruction time. Individual student time on Khan Academy also varied as some students were allowed to use the time dedicated to Khan Academy to complete paper worksheets assigned earlier in the week.

The teachers used Khan Academy as a supplement to their classroom instruction rather than as a primary source of instruction on new topics. Therefore, the Khan Academy resources most often used were the interactive problem sets and reporting dashboards; students' and teachers' use of the Khan Academy videos as an instructional aid was limited. Students tended to view the videos as a last recourse when they got stuck on practice problems. Typically, students used the "hints" feature first and then sought help from peers or their teachers before turning to the videos.

Students were often paired when using Khan Academy to promote peer-to-peer support. All Summit classrooms contain heterogeneous ability levels. During their use of Khan Academy during the school day, students worked on their own laptop and progressed at their own pace. Students were encouraged to seek help from their peers if they were struggling with a particular concept or problem within Khan Academy. The teacher was always present when students were using Khan Academy and walked the classroom monitoring student progress and working with students who requested assistance, both those working with Khan Academy and those who were engaged in other instructional activities.

The integration of students' use of Khan Academy into regular classroom curriculum progressed in two stages. In the first 4-6 weeks of the school year, there was no attempt to integrate Khan Academy into the curriculum. Students and teachers used this time as an introduction to Khan Academy and for students to work on filling gaps in basic skills as needed. All students started at the beginning of the Khan Academy "knowledge map" (e.g., simple addition) and worked through increasingly more difficult math topics at their own pace until most reached or went beyond grade-level expectations.

Approximately midway through the first semester, teachers began to integrate the use of Khan Academy with their classroom instruction by assigning weekly problem sets for students to complete. At the beginning of each week, teachers provided students with a list of problem sets that were associated with the content standards being covered during the whole-class instruction time and targeted in upcoming benchmark assessments. Teachers also used student performance on weekly in-class assessments to identify weekly Khan Academy goals for some students, asking those students who struggled on the weekly assessments to review and show mastery on the relevant content standards.

During November and December, the schools took a significant step in integrating Khan Academy into the curriculum. Khan Academy staff worked with teachers to develop an integrated blended unit on linear functions and systems of equations that used projects to help students apply new knowledge and skills. The units extended over 6 weeks. A member of Khan Academy's implementation team (a former Teacher for America teacher) mapped out the relevant videos and exercises for the unit for teachers and provided them with instructional strategies they could use to support self-paced instruction throughout the unit. Because the existing Khan Academy

[53] *A detailed description of the Summit's instructional and operational model is available in a case study developed as part of this research study and published on the Michael & Susan Dell Foundation website (www.msdf.org/blendedlearning).*

problem sets did not address the full set of discrete skills and problem types needed for an instructional unit on linear equations, Khan Academy staff developed additional problem sets for the unit.

The unit on linear equations consisted of three parts: daily teacher-directed instruction (whole-group, small-group, or individual instruction); independent work on Khan Academy to practice new skills; and a set of multistep problems that students worked on either independently or with a peer and allowed them to apply the skills they had practiced and mastered within Khan Academy.

The Khan Academy implementation coach worked with teachers to identify and develop the multistep applied problems. Two types of problems were developed: those that required independent self-directed exploration that usually took 30–40 minutes to complete, and those that were more comprehensive and required students to work with a partner who had made similar progress on the Khan Academy problem sets. Teachers developed a checklist that students used to guide them through the unit, including the sequence of activities that needed to be followed, the relevant Khan Academy problem sets, and the additional instructional resources students could use (including the relevant Khan Academy videos).[54]

On any given day during this unit, a variety of activities could be going on simultaneously in a classroom. For example, a teacher might be working with a small group of students or one-on-one with individual students in one part of the room, while other students worked independently on Khan Academy at the same time other students engaged in peer-tutoring or worked on a project independently or in pairs.

At the beginning of the second semester, the schools' math team kicked off collaboration with a local university that resulted in a different type of Khan Academy-supported unit. Teachers worked with the university faculty and three graduate students to develop a project-based geometry unit (Similarity, Trigonometry, and Congruence), a content area where Khan Academy had known gaps. Because the unit was designed around group work and whole-class pacing, initially there were few opportunities for self-paced instruction and the use of Khan Academy. Khan Academy staff worked with the teachers on a strategy to use Khan Academy videos and problem sets to support the unit, specifically to help students develop the prerequisite skills they would need to engage the project in meaningful ways. Students worked in groups of four throughout the unit, including during their time on Khan Academy. Each group was responsible for making sure all students in the group completed their Khan Academy problem sets, providing help to their peers when needed. For this particular unit, Khan Academy developed some new problem sets that were interactive and supported more conceptual understandings of the topic after observing teachers in the classroom teach the relevant topics. The teachers implemented the unit in February.

Sample

Table 1 summarizes some of the characteristics of the Summit schools included in the research. The two Summit schools that adopted a blended learning model, Tahoma and Rainier, enrolled a total of approximately 220 ninth-grade students in the 2011–12 school year. The two comparison schools, existing schools within the Summit Public Schools network, each enrolled approximately 300–400 students in Grades 9–12 (Everest had only Grades 9–11 at the time of the study). All of Summit's schools served an ethnically and economically diverse student population. Approximately 50% of the students qualified for the

TABLE 1. SUMMIT PUBLIC SCHOOLS IN RESEARCH SAMPLE

SCHOOL NAME (YEAR OPENED)	GRADE LEVELS 2011–12	STUDENTS (#)	2011–12 API RANKING[E]	2010–11 SIMILAR-SCHOOLS RANKING[F]
BLENDED LEARNING SCHOOLS				
RAINIER—SUMMIT SAN JOSE (2011)	9	110	848	NA
RAINIER—SUMMIT SAN JOSE (2011)	9	110	822	NA
COMPARISON SCHOOLS				
SUMMIT PREPARATORY CHARTER SCHOOL (2003)	9–12	440	825	9
EVEREST PUBLIC HIGH SCHOOL (2009)	9–11	330	829	10

[E] *The API ranges between 200 and 1,000, with 800 as the statewide goal for all schools.*

[F] *To create the Similar-Schools Ranking the California Department of Education compares a school's test scores to 100 schools across the state with similar demographic profiles. California uses parent education level, poverty level, student mobility, student ethnicity and other data to identify similar schools. Rank 1 means the school performed below at least 90 of its 100 similar schools. Rank 10 means the school performed above at least 90 of its 100 similar schools.*

[54] *A version of this model of Khan Academy use, including the use of checklists and self-paced instruction, were at the center of the schools' core math instruction during the 2012–13 school year.*

federal lunch program, and 45% were Hispanic. Academic Performance Index (API) scores in 2011–12 for Tahoma and Rainier were 848 and 822, respectively, while API scores for Everest and Summit Prep were 825 and 829, respectively. Similar-school rankings in 2010–11 for California Schools for the two comparison schools ranged between 9–10 or "well above average": these schools are at the top of the state academically relative to other public high schools serving similar student compositions (no rankings were available for the blended schools since they were new schools in 2011–12).

TABLE 2. DATA COLLECTION SUMMARY, FALL 2011 AND SPRING 2012 SITE VISITS, SUMMIT PUBLIC SCHOOLS

	FALL 2011 SITE VISIT	SPRING 2012 SITE VISITS
OBSERVATIONS		
Schools	3[A]	2[B]
Classrooms	5	3
INTERVIEWS		
CMO Leaders	3	2
School Leaders	4	3
Teachers	5	3
Students	NA	7

[A]. *Two San Jose schools and one of Summit's other high schools.*

[B]. *Two San Jose schools.*

SITE VISITS

The research team conducted site visits to Summit schools in fall 2011 and again in spring 2012. Each site visit was conducted over a single day. Table 2 summarizes the number of observation and interviews conducted during each site visit. For the fall 2011 site visit, the SRI team visited three Summit schools, including one of the two Summit high schools (Summit Prep) that was not using Khan Academy as an integral part of their math program. In the spring the research team returned to visit the two schools implementing Khan Academy.

TEACHER SURVEY

The research team conducted a survey of Summit's three teachers implementing a blended learning model using Khan Academy. All three teachers responded to the survey.

SYSTEM LOG DATA

System log data was provided by Khan Academy for students in Summit Public Schools' two San Jose schools. We examined both total time students spent actively engaged in Khan Academy instructional activities and the number of problem sets that a student successfully completed. Time spent in instructional activities covered students' viewing of videos and working on problem sets; reviewing reports or navigating other parts of the website did not contribute to our measure of total time of use of Khan Academy.

Students become "proficient" in a skill area (e.g., linear equations) when they correctly answer a certain number of questions within a problem set without making a mistake (typically 7 to 10 based on a machine-learning algorithm). For the time students spent working with Khan Academy instructional resources, we examined how time was distributed across viewing videos, working on problem sets, and between topic areas that were below or at or above the student's grade level (Grade 9), and how this variation might be explained by students' prior math ability. We then conducted a similar analysis for the number of problem sets successfully completed, examining how the topic areas of the problem sets completed varied by difficulty level (below versus at or above grade level), and how progress on Khan Academy problem sets varied across the sample as a whole and by prior ability. Finally, we examined the association between time-on-task and progress within Khan Academy and student learning outcomes.

STUDENT LEARNING OUTCOMES

Measures of learning outcomes included the California Standards Test (CST), Northwest Evaluation Association's (NWEA) Measurement of Academy Progress (MAP) math assessment, and an internal Algebra I benchmark assessment administered by Summit in all of its schools. The NWEA MAP is a computer-based adaptive assessment that is aligned with the California's Common Core State Standards (CCSS) (http://www.nwea.org). The NWEA MAP assessment was administered in the winter and spring within both San Jose schools and only one of the two comparison schools.[55]

[55] *An early fall 2011 administration of the NWEA MAP assessment was not possible in these sites since approval for the administration of the MAP and access to the test was not granted until several weeks into the school year.*

Implementation Findings

Findings from the site visits and teacher survey were used to help understand the facilitating factors and barriers to implementing the blended model during the 2011–12 school year. Additionally, the instructional system log data was used to help better understand the intensity with which students used Khan Academy in the blended learning model schools to supplement math instruction.

FINDINGS FROM THE SITE VISITS AND TEACHER SURVEY

Data from the fall 2011 and spring 2012 site visits and the teacher survey were used to report on different aspects of implementation of the blended model, including attempts to integrate Khan Academy into the curriculum, teachers' use of the system data to support their instruction, and teachers' training and support. In addition, we report on teachers' overall satisfaction with the blended model and the perceived benefits to students. A summary of the survey results is included in the Appendix for this profile.

SOFTWARE SELECTION

The fact that Khan Academy is an open education resource that was free to users and that Khan Academy agreed to support the schools and teachers at no cost to Summit were important factors in Summit leaders' decision to pilot Khan Academy. Although many aspects of Khan Academy appealed to Summit leaders, economy was a significant driver along with the Khan Academy's commitment to support the implementation by providing the schools with an implementation coach. Given restricted education budgets, Summit leaders were seeking cost-effective online instructional resources like Khan Academy to implement their instructional vision that did not require a significant financial investment in a single online instructional system.

INTEGRATION OF KHAN ACADEMY INTO THE CURRICULUM

Weekly goal setting was used to motivate students and align work in Khan Academy with the curriculum. To help motivate students and to align students' work on Khan Academy with the content being covered during regular instruction time and on interim benchmark assessments, teachers assigned up to 6 weekly problem sets within Khan Academy for students to complete. Goal setting was introduced approximately midway through the first semester, and by the second semester completion of the weekly goals became mandatory and contributed to a student's class grade. If the goals were not completed by the end of the day on Thursday, students were required to stay after school on Friday to complete them. If students completed the goals early, they were free to select more advanced goals that went beyond the curriculum.

Khan Academy was used to prepare students for state testing. At some point early in the second semester Khan Academy reorganized a series of problem sets of increasing difficulty into modules for teachers to use to prepare for end-of-year testing. The modules covered Arithmetic, Advanced Arithmetic, and Algebra. The modules were only made available to teachers participating in the pilot. Khan Academy recommended that teachers have students spend 15–20 minutes per day or 1–2 hours per week on the modules. Summit teachers reported using Khan Academy more often for test prep as the state-testing season approached. Teachers also supplemented the Khan Academy practice with other test prep resources.

TEACHER TRAINING AND SUPPORT

Khan Academy staff directly supported teachers in the classroom. Two of three teachers surveyed participated in two formal training sessions organized by Khan Academy during the first semester; the third teacher joined the school at the beginning of the second semester. Both teachers participating in training indicated they were very satisfied with the training they received. In addition to the formal training that was provided that introduced teachers to the instructional resources and the websites features, a Khan Academy implementation coach provided direct support to teachers throughout the school year. The implementation coach was frequently on campus observing classrooms and providing feedback and suggestions to teachers to help refine practices around the use of Khan Academy. In addition, staff from Khan Academy, including the implementation coach, participated in weekly meetings of the schools' math team to answer questions for the teachers, update teachers on new features and content being added to the website, and to collect feedback about the performance of the website and issues related to content coverage and student reporting.

Khan Academy viewed the use of its resources by Summit and other schools it partnered with in 2011–12 as a pilot and valuable source of user feedback that it used to develop and refine features on the website. Many changes were made to the Khan Academy platform, the content, and the organization of the content during the 2011–12 school year (SY). Starting in fall 2011, Khan Academy concentrated on building out its math content, mostly expanding the coverage of its exercises but also adding new videos where gaps in curriculum coverage were identified. By the end of the year, more than 3,000 YouTube videos and 350 problem sets were available. (At the start of SY 2011–12, there were approximately 2,500 videos and 130 problem sets.) Thirty-five problem sets were developed during January alone, including exercises on topics in arithmetic, algebra, statistics and probability, geometry, and complex numbers. Despite the near tripling of the number of problem sets available since the beginning of the school year, significant gaps still existed in the Khan Academy exercise topics relative to the CCSS. In some cases, these gaps limited teachers' ability to assign Khan Academy exercises linked to the curriculum topics they were covering. Khan Academy also added new features for students and teachers, including:

- The ability for students to set weekly goals for watching videos and for completing exercises
- New progress reports for students and teachers
- A search feature that enabled teachers to identify videos and exercises aligned with CCSS by grade level.

Khan Academy also modified its algorithm for assessing whether students were "proficient" in a particular skill area based on the number of problems answered correctly, moving from a fixed streak model (originally 10 correct answers in a row equated to proficiency) to a more nuanced model that used data from millions of users and machine-learning techniques to determine proficiency based on an estimated probability that a student will get the next problem correct given that the student correctly answered the previous problem (a probability of 94% was used).[56]

USE OF SYSTEM LOG DATA TO INFORM INSTRUCTION

Teachers' use of Khan Academy data evolved over the year with refinements made by Khan Academy in the reporting of student data. Early in the year, teachers made limited use of the data as the availability and format did not allow them to easily interpret and make timely instructional decisions for individual students. Toward the latter part of the first semester, Summit teachers and Khan Academy staff collaborated on the development of a reporting structure that enabled teachers to monitor individual student progress relative to their selected goals. Initially, the new reports created were made available to Summit teachers outside of the Khan Academy website through Google Docs. Later, these refinements to the student progress reports were integrated into the Khan Academy website and made available to all users.

SATISFACTION WITH BLENDED LEARNING AND BENEFITS TO STUDENTS AND TEACHING

Summit teachers were generally positive in their responses about the school's blended learning model, and all three would recommend the use of Khan Academy to other teachers. Teachers indicated that Khan Academy was at least somewhat effective for students whose academic work was ahead, at the expected level, or behind most students their age. Two of the three teachers reported that Khan Academy was "very effective" with students who were performing below their grade level peers. However, all three teachers also responded in the survey that a lack of alignment between Khan Academy content and the school curriculum, most likely due to gaps in the Khan Academy content for ninth grade, had a "moderate" or "significant" impact on their ability to use Khan Academy effectively with their students. Khan Academy continues to work on building out its high school content.[57]

Khan Academy was used to support the development of procedural skills and facilitate one-on-one and small-group instruction. In a unit on integrated linear equations, teachers used Khan Academy in a variety of ways to support their instruction to help students develop the procedural skills needed to engage in multistep complex problems and activities that required higher-order thinking. Students were not allowed to work on these activities until they successfully completed the relevant Khan Academy problem sets. The teachers reported that using Khan Academy to provide students with self-paced instruction on procedural skills also freed them up to work with individual and small groups of students who were having the greatest difficulties mastering the relevant procedural skills and to support students who had successfully moved onto the applied problem activities.

[56] *For details, see http://david-hu.com/2011/11/02/how-khan-academy-is-using-machine-learning-to-assess-student-mastery.html*

[57] *One of Summit's teachers participating in the research worked with Khan Academy staff during summer 2012 to identify gaps in the Grades 9–12 content and develop new content.*

Self-paced instruction required teachers to provide custom support to students at different places in the curriculum. Because many of the applied problems in the linear equations unit were self-paced, the time it took students to complete the unit varied widely. According to the Khan Academy implementation coach and the teachers, some students finished with 10 days left in the unit, while other students needed extra time beyond the scheduled end of the unit. Students who finished the unit early moved on to other activities including tutoring their peers, working on extension activities provided by the teacher, and working on SAT-prep exercises.

The teachers reported that they were initially unprepared to support students who progressed through the unit quicker than their peers. While teachers wanted to be able to challenge these accelerated learners, teachers did not want them to continue in Khan Academy and get too far ahead of the rest of the class since the teachers had not planned out activities to supplement the Khan Academy instruction. Teachers reported they often supplemented Khan Academy for these students with other resources on linear equations since there wasn't enough content available within Khan Academy to support extension activities that would allow these students to go deeper into the topic area.

Teachers also supplemented Khan Academy for struggling learners. Teachers reported that they frequently supplemented Khan Academy resources with other online resources for those students who struggled with the Khan Academy problem sets. The teachers expressed that in some ways they favored supplementing the Khan Academy instruction with other online resources for these students since it provided students with alternative resources to draw upon and broke up the tedium of working with the same online resources day after day. The median student in Summit worked on Khan Academy for 102 days of the school year.

Student engagement varied by ability level. Teachers reported during the interviews that they felt that engagement levels throughout the year were consistently high for high- and middle-achieving students. All three teachers felt that high-achieving students liked working with Khan Academy as it allowed them to explore new concepts in math that went beyond the curriculum. One teacher felt that the students who were most engaged when using Khan Academy in his classroom were a group of students who excelled in their other coursework but not in math. For these students, Khan Academy gave them an opportunity to work independently at their own pace until they had some success. A sense of progress with Khan Academy, something that they were not experiencing in the regular math classroom, was extremely motivating for these students and pushed them to work harder than they might have otherwise, according to the teacher. All three teachers were in agreement that most students at the lower end of the achievement distribution, who struggled in the classroom, also struggled to make progress within Khan Academy, and this constant struggle was demoralizing. One teacher commented that these students may have had a different experience with Khan Academy if they were allowed to progress through the content at the appropriate level and at their own pace. However, after the first 8 weeks of the school year, all students in a classroom were expected to complete the same teacher-assigned problem sets by the end of the week that were linked to the topics the teacher covered during regular instruction. As result, these students were often forced to confront concepts and skills that were beyond their current preparation.

FINDINGS FROM THE ANALYSIS OF INSTRUCTIONAL SYSTEM LOG DATA

To capture the intensity with which students used Khan Academy and the progress made within Khan Academy in the two Summit schools, we computed both the total number of videos students watched and problem sets successfully completed over the school year, as well as the total number of days and minutes that students interacted with videos and problem sets. In addition, we computed the amount of problem sets and time spent interacting with problem sets that were below grade level or at or above grade level.[58] Table 3 provides the summary statistics for these Khan Academy time-on-task and progress indicators.

Students spent a considerable amount of time working on problem sets. Up to 45 minutes per day, 4 days per week, was allocated by the schools for use of the Khan Academy resources across the school year. Some students also used Khan Academy outside of school hours, on their own and to complete weekly assignments. The median student spent 3,422 minutes (or slightly less than

[58] *A complete mapping of video content to grade-level expectations was not available at the time of the analysis.*

TABLE 3. DESCRIPTIVE STATISTICS FOR INTENSITY OF USE AND PROGRESS INDICATORS	MEAN (MEDIAN)	STANDARD DEVIATION	MINIMUM	MINIMUM
Number of Days[a]	100 (102)	19	27	223
Minutes (Total)	3,422 (3,140)	1,503	123	10,167
Video Minutes	132 (109)	143	3	1,640
Videos Watched	29 (23)	30	1	384
Exercise Minutes	3,169 (2,869)	1,431	144	8,864
Below Grade Level	1,421 (1,366)	588	17	3,599
At/Above Grade Level	1,735 (1,493)	929	126	5,265
Problem Sets Completed	168 (167)	51	2	361
Below Grade Level	114 (111)	35	1	216
At/Above Grade Level	54 (56)	18	1	144

[A] *based on smaller sample (N = 149) due to missing data.*

90 minutes per week for a 36 week school year) working on Khan Academy (either on exercises or viewing videos) during the 2011–12 school year. The median number of days used was 102 (see Table 3). Approximately 90% of students' time was spent working on problem sets (about half of which was content below the ninth-grade level), while the rest of time was spent viewing videos.

The average student viewed 29 videos across the entire school year. Since Khan Academy was used primarily as a supplement to the core curriculum, the videos were not assigned as a primary source of instruction. Teachers were still responsible for providing initial instruction on key concepts covered by the curriculum, typically in the form of whole-class lectures with guided practice activities. Khan Academy was used to provide opportunities for students to practice and master recently learned skills. While working on problem sets, students rarely relied on the videos for support, often times as a last resort after several unsuccessful attempts at solving a set of problems in a particular topic area. When students struggled with a particular exercise, they were more likely to rely on support from the "hints" option built into each question or the assistance of their teacher or peers.

A majority of the content worked on was below grade level. The median number of problem sets successfully completed was 167 (or more than 4 per week). More than 65% of the problem sets completed were on content below the ninth-grade level, with the other 35% at or above grade level.[59] The discrepancy between the grade levels of the exercises completed was likely a result of how the Khan Academy was used during the first half of the first semester compared to the rest of the year. At the start of the school year, as part of an orientation to the Khan Academy resources, students were instructed to explore the resources, at their own pace, by starting with the lowest level content available (e.g., simple arithmetic) and then encouraged to work their way up to grade level content, and beyond if possible. During this time teachers also assigned about an hour of Khan Academy work to be completed outside of the classroom. By the end of October, teachers started implementing weekly goal setting, and eventually began to assign weekly problem sets (4 to 6 per week) that were linked to the current topics being covered in the curriculum and that students were required to complete as part of their class grade. As a result, it is likely that a majority of the below grade-level problem sets completed were completed during the first half of the first semester.

Time spent working on Khan Academy and progress made varied by prior academic performance (Figures 1–2 and Table 4). There were significant differences between students with different prior achievement levels in the amount of time students spent working on Khan Academy and the number of problem sets successfully completed. Students who scored below the median on a math diagnostic assessment administered by the schools prior to the school year spent about 12% more time working with Khan Academy resources (about 400 more minutes or more than 6 hours) than

[59] *For high school, only below and at grade level (9–12) classifications of content were available. As a result, all problem sets mapped to Grades 9–12 are indicated as being at grade level for the Grade 9 students in Summit.*

Michael & Susan Dell FOUNDATION

TABLE 4. RELATIONSHIPS BETWEEN PRIOR ACHIEVEMENT LEVELS AND KHAN ACADEMY USE INDICATORS FOR SUMMIT PUBLIC SCHOOLS

USE INDICATOR	LOWER PRIOR	HIGHER PRIOR
Number of Days Used[A]	102	103
Total Minutes	3,759	3,362*
Video Minutes	125	150
Videos Watched	28	32
Problems Attempted	3,926	5,118*
Exercise Minutes	3,533	3,058*
Below Grade Level	1,654	1,306*
At/Above Grade Level	1,868	1,738
Problem Sets Completed	148	204*
Below Grade Level	101	138*
At/Above Grade Level	47	66*

[A] *Note: N = 180*

a based on smaller sample (N = 149) due to missing data.

Significance tests assessed differences between prior performance groups (Lower prior and Higher prior) and were run using a 2-Level hierarchical linear model where students were nested within schools.

*$p < .05$

students who scored above the median on the same diagnostic assessment (see Table 4). Although these students spent more time working with Khan Academy, students with higher prior achievement completed more problem sets both below and at and above grade level (almost 30% more or on average 40 additional problem sets). These differences in minutes used and progress made by prior achievement are consistent within the self-paced and personalized instructional environment adopted by Summit. Students who needed more time to master the content could do so at their own pace, while more advanced students demonstrated mastery on the assigned problem sets more quickly and were able to explore new topic areas that were units or grade levels ahead of the curriculum.

Figures 1 and 2 further illustrate how time-on-task within Khan Academy and problem sets completed varied by prior achievement. To produce these figures, students were ranked ordered on their

FIGURE 1. AVERAGE TOTAL MINUTES WORKING ON KHAN ACADEMY PROBLEM SETS BELOW AND AT OR ABOVE GRADE LEVEL BY PRIOR ACHIEVEMENT

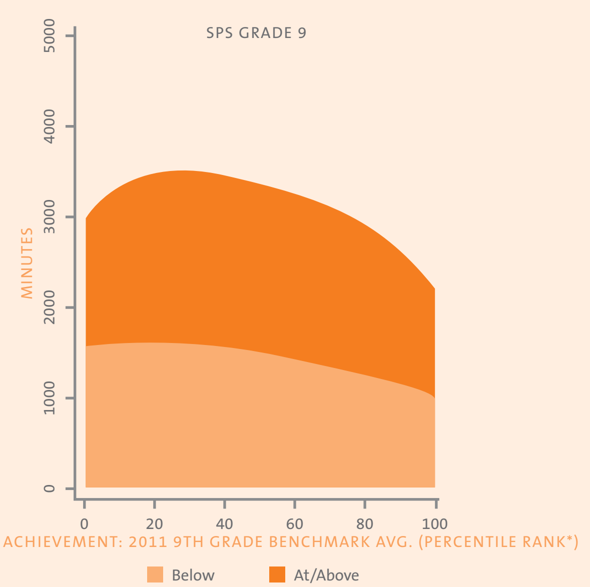
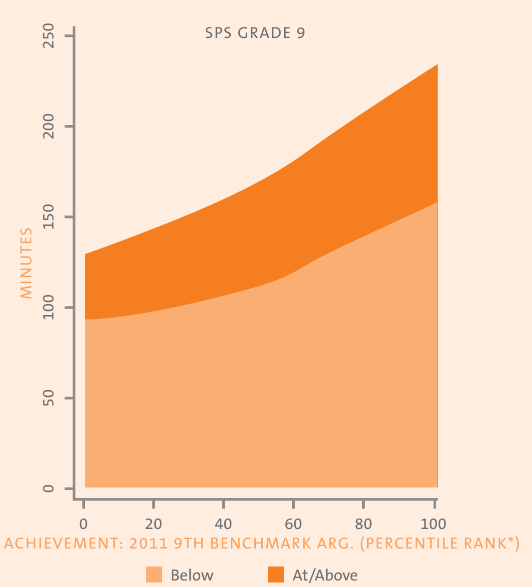


FIGURE 2. AVERAGE TOTAL NUMBER OF PROBLEM SETS COMPLETED BELOW AND AT OR ABOVE GRADE LEVEL BY PRIOR ACHIEVEMENT.



prior achievement scores, and then these rankings were mapped to the time they logged viewing videos and working on problem sets, as well as the number of problem sets successfully completed. The colored bands on these exhibits represents the extent to which students with similar prior achievement scores spent time working on Khan Academy and successfully completed problem sets on content that was below (light orange) or at or above grade level (orange). Figure 1 illustrates how the number of minutes using Khan Academy decreased with increasing prior achievement scores, while Figure 2 shows how the number of problem sets successfully completed increased for content both below and at or above grade level.

Impact Analyses

IMPACT DESIGN OVERVIEW

A quasi-experimental design was implemented to explore the relationship between the adoption of Khan Academy as a supplement to support math instruction and gains in student learning. The design involved comparing the learning outcomes for students enrolled in Summit's San Jose schools to comparison students attending the two other schools in the Summit Public School charter school network where use of Khan Academy, when it was used, was limited and sporadic. Both the San Jose schools and the comparison schools serve diverse student populations. Scores for spring learning outcome assessments for students in the San Jose and the comparison schools were compared.[60] Table A.1 in the Appendix for this profile shows the baseline descriptive statistics for students included in the sample for the impact analysis. A few differences between the Summit's San Jose school and comparison schools should be noted. The San Jose schools had more than twice as many students who were Asian (20% compared to 9.5%) and half as many students who were designated as White (17% compared to 35%). The San Jose schools also had more than twice as many students who were eligible for the federal subsidized lunch program (45% compared to 21%). Despite these differences between the San Jose schools and the comparisons in the composition of the students, there was no statistically significant difference in prior achievement as measured by an internal assessment of ninth-grade Algebra-readiness.

Limitations. Although the comparison schools may be similar in many ways to the schools using Khan Academy, there may have been important differences between the San Jose and the comparison schools (e.g., differences in curriculum, academic culture, school leadership, and teaching staff) that may explain differences between the schools on spring test scores that are totally unrelated to the use of Khan Academy. Thus, using this design we could not isolate the effect of the introduction of Khan Academy from other key differences between schools that are likely to influence student achievement. For example, in the San Jose schools, all ninth graders were enrolled in a single, year-long integrated Algebra and Geometry course supplemented by the use of Khan Academy. There were three teachers that taught the course across the two schools. In the comparison schools, most ninth graders, but not all, took separate Algebra I and Geometry I courses taught by four different teachers, two different teachers in each school.

IMPACT FINDINGS

Main impacts (Tables 5 and 6). Results for Summit schools were mixed; on the state's test of standardized achievement (CST), students in schools using Khan Academy outperformed Summit schools that were not, but on Summit's own internal Algebra I benchmark assessment, no differences were found.[61] Specifically, students in Grade 9 in the two San Jose schools outperformed their peers in Summit's other two high schools on the CST (effect size

[60] *Spring learning assessment scores were available for CST and benchmark assessments for both the San Jose and the comparison schools. Scores for the NWEA MAP assessment were only available for one of the two comparison schools and so the scores for the NWEA MAP were not included in the impact analysis.*

[61] *We originally planned to also analyze the impact on the NWEA MAP test scores. However, only one of the two comparison schools chose to implement the additional assessment. Summit's schools were implementing the MAP for the first time for the purpose of the research. As a result, school and teacher buy-in into the value of the assessment for informing instruction varied across the schools and implementation of the winter and spring MAP was uneven, resulting in unreliable growth estimates for participating schools in these sites. In addition, all schools experienced significant technical issues with the implementation of NWEA's online version of the test that disrupted test administration (particularly during the winter test window) and contributed to a less than ideal testing experience for schools, teachers, and students.*

TABLE 5. EFFECT SIZES BY OUTCOME MEASURE AND SUBGROUP

			SUBGROUP							
			GENDER		PRIOR ACHIEVEMENT		ENGLISH LANGUAGE LEARNER (ELL)		FREE OR REDUCED-PRICE LUNCH (FRPL)	
	GRADE LEVEL	MAIN EFFECTS	F	M	ABOVE MEDIAN	BELOW MEDIAN	ELL	NOT ELL	FRPL	NOT FRPL
SUMMIT										
Benchmark	9	-0.05	-0.14	-0.09	0.29*!	-0.42*!	0.36!	-0.19!	-0.04	-0.11
CTS	9	0.54*	0.43*	0.43*	0.91*!	0.10!	0.19	0.49*	0.25	0.56*

** Contrast between treatment and control significant for this subgroup, p < .05*
! Treatment effect significantly varied by subgroup, p < .05

TABLE 6. EFFECT SIZES AND CORRESPONDING MEAN SPRING CALIFORNIA STANDARDS TEST (CST) TEST SCORES AND CST PERFORMANCE LEVELS BY EXPERIMENTAL GROUP

	ADJUSTED MEAN CST SCORE				KHAN ACADEMY (% AT PERFORMANCE LEVELS)					COMPARISON(% AT PERFORMANCE LEVELS)				
	EFFECT SIZE	KHAN ACADEMY	COMPARISON	DIFFERENCE (CTS SCALE POINTS)	FBB	BB	B	P	A	FBB	BB	B	P	A
Summit (9th)	+.54*	402.2	363.0	39.2	3	15	20	33	29	5	10	29	38	18

Performance levels: FBB = Far Below Basic, BB = Below Basic, B = Basic, P = Proficient, A = Advanced
Contrast between Treatment and Comparison statistically significant, p < .05
Note: The mean scores and differences are "adjusted" model-based estimates. However, for the descriptive percentages for performance levels there are no statistical controls for prior achievement.

= +.54; statistically significant). On the CST, the effect of being in a school using Khan Academy was associated with an improvement of 39 scale points (402 compared to 363) or reflects the median student in the comparison schools improving 18 percentiles if that student had enrolled in one of the San Jose schools using Khan Academy at the start of the 2011–12 school year (i.e., the median student in the comparison schools moving from the 50th to 68th percentile). When CST performance levels are examined between students in the San Jose and comparison schools (Table 6), we see a difference favoring the San Jose schools of 6 percentage points in the number of students achieving at the proficient or advanced proficient (62% compared with 56%).[62]

Subgroup analyses (Table 5). We found significant differences in the estimated effects for students with different levels of prior achievement. On the CST and Algebra I benchmark, students in the San Jose schools who scored at or above the median on a math placement assessment administered within all Summit schools before the school year outperformed students who scored below the median on the placement test relative to similar student in the comparison schools. On the Algebra benchmark assessment, where for the overall sample we found no difference in how students scored between the two groups of schools, when we disaggregated the effects by prior achievement levels we found a positive and statistically significant effect for students with higher prior achievement levels and a negative and statistically significant effect for students with lower prior achievement. When examining CST scores, we found a somewhat similar result, again favoring students with higher prior achievement. Specifically, students with higher prior achievement in Summit's San Jose schools significantly

[62] *In an attempt to rule out the possibility something other than the schools' math instructional model was responsible for the effects on math performance, we also analyzed the students' performance on the ELA CST, a measure unrelated to the use of Khan Academy or math instruction. We found no effect of attending the San Jose schools on students' performance on the ELA CST. Thus, it does not appear that differences in the types of students attending the different schools or some difference in a global school factor (e.g., academic culture) is responsible for the differences in math performance between the groups; otherwise, we might expect to see similar gains in all measures of achievement.*

TABLE 7. RESULTS OF ANALYSIS OF KHAN ACADEMY USE PREDICTING EXPECTED STUDENT ACHIEVEMENT, CST SCORES (SUMMIT PUBLIC SCHOOLS)

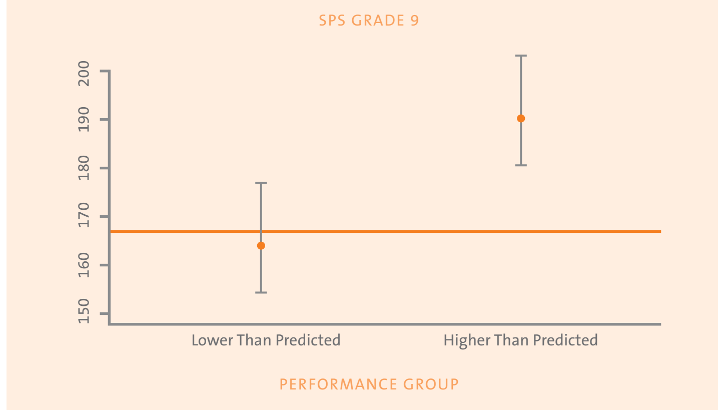
	LOWER THAN PREDICTED			HIGHER THAN PREDICTED		
	N	MEAN	SD	N	MEAN	SD
Khan Academy Minutes	76	3,584	1,515	67	3,626	1,532
Problem Sets Completed ***	76	165	45	67	190	47

Note: Significance tests assessed differences between performance groups (Lower than predicted and Higher than predicted) and were run using a 2-Level hierarchical linear model where students were nested within schools.

SD = standard deviation

**** = p < .001*

FIGURE 3. PROBLEM SETS SUCCESSFULLY COMPLETED BY LOWER AND HIGHER THAN PREDICTED STUDENT GROUPINGS.



outperformed their peers in the comparison schools on the CST, while there was no difference in the spring test scores between students with lower prior achievement in the two groups. In addition, the effects on scores on the Algebra I benchmark varied by English language status. Specifically, we found a positive effect for students designated as English language learners and a negative effect for students who were not designated as such. The differences in the effects were statistically significant. We found no statistically significant difference in the effects between students who were or were not eligible for the free or reduced-price lunch program.

EXPLORING THE LINK BETWEEN USE AND STUDENT OUTCOMES

The relationships between time-on-task, progress, and scores on the spring CST and NWEA MAP assessment were examined within the two Summit schools using Khan Academy. All analyses controlled for prior student achievement, an important consideration given that use, progress, and prior student ability were positively correlated. Students were grouped by whether their spring test scores exceeded or were lower than what was predicted based on their prior academic performance (score on an internal math placement test). We then compared the average number of minutes for time-on-task and problem sets completed for the two groups to see whether more time on Khan Academy or greater progress was associated with better than expected spring achievement scores. While we attempted to control for prior achievement differences between students in these analyses, by design these analyses are exploratory; the results cannot be used to make definitive claims about the effectiveness or noneffectiveness of the use of Khan Academy resources to supplement core math instruction. Instead, the results should be used to suggest possible causal relationships and areas of future research.

Within the two Summit schools that used Khan Academy as an integral part of the math curriculum in 2011–12, we found a statistically significant difference for the numbers of problem sets completed by the two groups of students, but not for time-on-task. Results for the California Standards Test (CST) are shown in Table 7 and Figure 3 (the results using the NWEA MAP assessment were similar). Those students who had higher than predicted spring CST and MAP scores on average completed more Khan Academy exercises during the course of the year than students whose scores were lower than predicted, about 20% more (or 35 exercises). The differences between the groups were statistically significant.

Michael & Susan Dell FOUNDATION

Appendix: Summit Public Schools

TABLE A.1. DESCRIPTION OF SUMMIT'S SAN JOSE AND COMPARISON STUDENTS IN THE ANALYTIC SAMPLE

VARIABLE	SAN JOSE		COMPARISON	
	%	N	%	N
GENDER				
Male	48.6	102	50.2	110
Female	51.4	108	49.8	109
ETHNICITY				
American Indian	1.9	4	0.0	0
Asian	20.0	42	9.5	21
Black or African American	6.2	13	5.0	11
Hispanic or Latino	54.8	115	47.1	104
Pacific Islander	0.0		2.7	6
White	17.1	36	35.7	79
FREE/REDUCED-PRICE LUNCH				
Eligible	45.4	93	21.4	47
Not eligible	54.6	112	78.6	173
ENGLISH LANGUAGE LEARNER STATUS				
ELL	18.5	38	16.4	36
Non-ELL	81.5	167	83.6	184
PRIOR ACHIEVEMENT	MEAN (STANDARD DEVIATION)	N	MEAN (STANDARD DEVIATION)	N
9th-Grade Algebra-Readiness[a]	40.9 (22.5)	210	43.8 (24.9)	219

[A] *No statistically significant differences between San Jose and comparison schools on scores of ninth-grade algebra-readiness (effect size = -.12, not significant at the p < .05 level) were found.*

TABLE S1. TEACHER SURVEY RESPONDENT CHARACTERISTICS

TEACHER CHARACTERISTCS	NUMBER	PERCENTAGE
SUBJECT TAUGHT		
Math	3	100
TEACHING EXPERIENCE		
New teachers (3 years or less)	2	66
Veteran teachers (4 or more years)	1	33

Michael & Susan Dell FOUNDATION

TABLE S2: COMPARISON OF NUMBER OF TEACHERS SPENDING DIFFERENT LEVELS OF TIME IN INSTRUCTIONAL ACTIVITIES BY BLENDED LEARNING AND COMPARISON SCHOOLS

IN A TYPICAL MATH CLASSROOM SESSION, HOW MUCH TIME IS SPENT IN THE FOLLOWING ACTIVITIES?	BLENDED COUNTS				COMPARISON COUNTS			
	0 MIN.	1-15 MIN.	16-30 MIN.	OVER 30 MIN.	0 MIN.	1-15 MIN.	16-30 MIN.	OVER 30 MIN.
Teacher-led whole-class instruction	0	2	1	0	0	2	0	0
Teacher-led small-group instruction	0	3	0	0	0	2	0	0
Teacher-led one-on-one instruction	0	2	1	0	0	2	0	0
Small-group collaborative projects	0	0	2	1	0	0	2	0
Students engaged in independent work/practice	0	0	0	3	0	2	0	0
Students engaged in self-directed instructional activities	0	1	1	1	0	2	0	0

TABLE S3. NUMBER OF TEACHERS REPORTING DIFFERENT ROLES FOR TECHNOLOGY AND WEB-BASED INSTRUCTION IN SUPPORTING TEACHER-PROVIDED INSTRUCTION

WHAT ROLE DOES TECHNOLOGY AND WEB-BASED INSTRUCTION PLAY IN SUPPORTING THE INSTRUCTION YOU PROVIDE?	MAJOR ROLE	MODERATE ROLE	MINOR ROLE	NO ROLE
To provide students with practice exercises during the school day	3	0	0	0
To capture student data to inform instruction	2	1	0	0
For homework	2	1	0	0
To allow students to learn at their own pace	2	1	0	0
For remediation/to fill in gaps in knowledge	1	2	0	0
For enrichment for advanced students	1	2	0	0
To introduce new concepts within a particular lesson (i.e.,embedded within delivery of core lesson)	1	1	1	0
For diagnostic or formative assessment	1	0	2	0
To meet needs/interests of different types of learners	1	0	2	0
To provide an additional way to access material	0	2	1	0
To help facilitate small-group face-to-face instruction	0	1	2	0
For test prep for standardized or state assessments	0	1	2	0
To promote deeper learning (e.g., critical thinking, collaboration)	0	1	2	0

Teacher Survey Results – Summit Public Schools

SURVEY ADMINISTRATION

Overall, 3 of 3 Summit teachers implementing blended learning completed the Teacher Survey for a 100% response rate. Table S1 describes the characteristics of the respondents in terms of the subject area(s) they teach and years of teaching experience. Additionally, 2 of 2 teachers at nonblended schools completed the Comparison Teacher Survey for a 100% response rate.

CLASSROOM INSTRUCTIONAL ACTIVITIES

Although our survey information is based on small numbers, we found more emphasis on students engaging in independent work/practice and self-directed instructional activities in the blended schools than in the comparison nonblended schools; reports of instructional time devoted to the other activities listed in Table S2 were similar across the blended and comparison schools.

Michael & Susan Dell FOUNDATION

TABLE S4. NUMBER OF TEACHERS USING DATA "SOMEWHAT" OR "A GREAT DEAL" FOR DIFFERENT INSTRUCTIONAL PURPOSES

TO WHAT EXTENT HAS YOUR REVIEW OF STUDENT PERFORMANCE DATA FROM THE SYSTEM(S) LED YOU TO DO THE FOLLOWING?	"A GREAT DEAL" OR "SOMEWHAT" (COUNT)
Identify gaps in student learning or comprehension	3
Modify plans for future lessons or instructional activities	3
Modify topics covered on future student assessments (e.g., to confirm patterns found in student performance data)	2
Group or identify students for pull-outs	3
Provide feedback to individual and/or small groups of students	3
Monitor and diagnose individual student's understanding of key concepts	3
Monitor and diagnose whole class or a group of students' understanding of key concepts	3
Set expectations / goals for student achievement	2

ROLE OF TECHNOLOGY

All three teachers indicated that technology and web-based instruction played a "major" role in supporting their instruction by providing students with practice exercises during the school day. The role technology and web-based instruction played in introducing new concepts within a particular lesson varied among Summit teachers, ranging from a minor to major role (see Table S3).

USE OF DATA

All (100%) of Summit's teachers responding to the survey indicated they reviewed the student performance data recorded by the software system(s) used to support blended learning; two teachers indicated they view the data every day or almost every day, and one teacher indicated he or she viewed it two or three times a week. All (100%) reported the student performance data captured by Khan Academy were "somewhat" useful in informing their instruction. They used the data "somewhat" or "a great deal" for a variety of purposes, including for guiding whole-class instruction and for supporting individual and small groups of students (see Table S4).

Satisfaction with Blended Learning and Impact on Student Learning

TEACHER SATISFACTION WITH BLENDED LEARNING

Summit teachers were generally positive in their responses about the schools' blended learning model. For example:

- 2 of 3 agreed Khan Academy met the learning needs of their students
- 3 of 3 agreed that students were highly engaged while using Khan Academy
- 2 of 3 agreed that Khan Academy helped students take ownership for their own learning
- 2 of 3 agreed that students' learning and understanding of the material improved due to the use of Khan Academy
- 3 of 3 agreed they would recommend the use of Khan Academy to other teachers

Michael & Susan Dell FOUNDATION

EFFECTIVENESS FOR STUDENT TYPES

Teachers indicated that Khan Academy was at least somewhat effective for students whose academic work is ahead, at the expected level, or behind most students their age. In fact, two of three teachers felt it was "very effective" with students who were performing below their grade-level peers. The teachers were divided in their opinions regarding the effectiveness for special education students, with one indicating Khan Academy was very effective, another saying it was not at all effective, and a third unsure about the effectiveness.

EFFECTIVENESS RELATIVE TO NONBLENDED ENVIRONMENT

One of the three Summit teachers completing the survey had 4 or more years of teaching experience. This individual indicated that, compared to a traditional nonblended environment, students did better in a variety of skill areas with blended learning, including recalling facts/terms/basic concepts, demonstrating comprehension, applying knowledge in new situations, and demonstrating higher-order thinking (e.g., analyze, synthesize, evaluate). The two new teachers differed in their opinions, with one indicating they did not know and the other indicating a positive impact in three of the four skills areas, and a negative impact related to high-order thinking skills.

TRAINING AND SUPPORT

Two of three teachers participated in a training or orientation session (either in person or online) directly related to their school's use of blended learning or on the specific software program(s) supporting it; the third indicated no training was provided. Both teachers participating in training indicated they were very satisfied with the training they received to use Khan Academy. The amount of their own time teachers reported spending to get acquainted with Khan Academy or planning for how to best integrate blended learning with their instruction varied, with one teacher spending less than an hour and another spending as many as 11–20 hours.

FACTORS INFLUENCING USE

All three teachers cited a lack of alignment between computer-based instruction and the school curriculum as having a "moderate" or "significant" impact on their or their school's ability to use Khan Academy effectively. Two of three also cited a lack of access to technology for students.

Although all (100%) the teachers reported that they or their students experienced technical problems in using Khan Academy, only one indicated that technical challenges were any more than a "slight" barrier to effective use of Khan Academy. The most commonly cited technical problems included student computers not working properly (3 of 3), Internet connection unreliable/down (3 of 3), insufficient bandwidth/Internet too slow (2 of 3), and software glitches or bugs (2 of 3).

Technical Appendix

This appendix describes the preparation and modeling of academic achievement data from the five participating sites. The first section describes the overall process for collecting data, cleaning and preparing specific variables for use in statistical models. The second section describes the statistical models used to estimate impacts (including impacts by subgroups). Lastly, a set of reference tables appears in the third section showing the model parameters resulting from the different sets of analyses.

DATA PREPARATION

Demographic and academic achievement data were provided directly from the five participating sites. At the beginning of the research project the sites were given Microsoft Excel data templates outlining the specific variables that would be needed for the analyses. Sites varied somewhat in their strict compliance with the template format; nonetheless all sites provided the requisite data elements. In addition to these data templates, sites using the NWEA MAP assessment delivered the score files for MAP examinees.

In order to maintain confidentiality of student achievement records, sites assigned ID numbers to each student before compiling the research data. The data sets were then deposited to a specific dedicated secure file server to which a limited number of SRI staff had access. A data analyst not otherwise involved with the project would then substitute a consistently formatted SRI-generated ID number for the site specific IDs before releasing the file to the analysis team.

The data files were merged as needed and data elements checked for appropriate values. The records themselves were checked for duplicate ID values in the cases where none should have been present. Where questions or discrepancies appeared, the sites were contacted to resolve the issues.

Not all requested data elements were provided by every site, and in some cases there was a significant degree of missing data for particular elements. Table A.1 below describes the number of usable data elements by site, grade level, and (as appropriate) year. A key for these tables is shown below.

ELL = English language learner status
FRPL = Federal subsidized lunch program status
IEP = Individualize Education Program (special education status)
ELA = English/Language Arts
Benchmark = Internal achievement assessments administered by the sites
MAP = Northwest Evaluation Association Measure of Academic Progress
CST = California Standards Test (California sites only)
i/LEAP = Louisiana Educational Assessment Program assessment (FirstLine Schools only)

Michael & Susan Dell FOUNDATION

Table F.1. Number of data elements (and percent of cases) by site, grade, and year

Alliance College-Ready High Schools

	2010		2011		2012	
Alliance	N	%	N	%	N	%
9TH GRADE						
ELL					656	73%
FRPL					790	88%
Gender					464	52%
Benchmark (ELA)					874	97%
CST ELA					852	95%
CST Math					525	59%
MAP Math					510	57%
MAP ELA					454	51%
GRADE 10						
ELL					579	73%
FRPL					694	88%
Gender					405	51%
Benchmark (ELA)					767	97%
CST ELA					0	0%
CST Math					0	0%
MAP Math					410	52%
MAP ELA					395	50%

Michael & Susan Dell FOUNDATION

Firstline Schools

	2010		2011		2012	
Firstline	N	%	N	%	N	%
GRADE K						
ELL	0	0%	18	9%	241	97%
FRPL	0	0%	18	9%	242	98%
Gender	82	60%	108	52%	133	54%
IEP	0	0%	15	7%	230	93%
Benchmark (ELA)	0	0%	0	0%	1	0%
Benchmark (Math)	0	0%	0	0%	0	0%
I/Leap ELA	0	0%	0	0%	0	0%
I/Leap Math	0	0%	0	0%	0	0%
TerraNova Math	137	100%	208	100%	227	92%
TerraNova Reading	137	100%	208	100%	230	93%
GRADE 1						
ELL	0	0%	34	16%	228	98%
FRPL	0	0%	34	16%	224	96%
Gender	77	52%	124	57%	123	53%
IEP	0	0%	27	13%	198	85%
Benchmark (ELA)	0	0%	55	25%	0	0%
Benchmark (Math)	0	0%	56	26%	0	0%
I/Leap ELA	0	0%	0	0%	0	0%
I/Leap Math	0	0%	0	0%	0	0%
TerraNova Math	148	100%	205	95%	221	95%
TerraNova Reading	148	100%	205	95%	221	95%
GRADE 2						
ELL	0	0%	34	15%	211	98%
FRPL	0	0%	32	14%	213	99%
Gender	82	51%	122	53%	125	58%
IEP	0	0%	24	10%	185	86%
Benchmark (ELA)	0	0%	69	30%	209	97%
Benchmark (Math)	0	0%	69	30%	210	97%
I/Leap ELA	6	4%	7	3%	0	0%
I/Leap Math	6	4%	7	3%	0	0%
TerraNova Math	155	96%	218	94%	204	94%
TerraNova Reading	154	96%	217	94%	203	94%
GRADE 3						
ELL	17	20%	50	22%	246	98%
FRPL	0	0%	45	19%	243	97%
Gender	42	50%	118	51%	131	52%
IEP	14	17%	37	16%	210	84%
Benchmark (ELA)	0	0%	231	100%	240	96%
Benchmark (Math)	0	0%	230	99%	241	96%
I/Leap ELA	55	65%	224	97%	233	93%

Michael & Susan Dell FOUNDATION

Firstline Schools (cont.)

	2010		2011		2012	
Firstline	N	%	N	%	N	%
GRADE 3 (CONT.)						
I/Leap Math	55	65%	224	97%	233	93%
TerraNova Math	37	44%	109	47%	230	92%
TerraNova Reading	37	44%	109	47%	232	93%
GRADE 4						
ELL	26	42%	52	25%	261	100%
FRPL	0	0%	51	24%	250	95%
Gender	36	58%	107	50%	136	52%
IEP	19	31%	39	18%	177	68%
Benchmark (ELA)	0	0%	209	99%	240	92%
Benchmark (Math)	0	0%	207	98%	244	93%
I/Leap ELA	53	85%	202	95%	214	82%
I/Leap Math	53	85%	201	95%	214	82%
TerraNova Math	16	26%	41	19%	0	0%
TerraNova Reading	16	26%	41	19%	0	0%
GRADE 5						
ELL	22	30%	24	13%	229	98%
FRPL	0	0%	23	12%	229	98%
Gender	42	57%	100	54%	117	50%
IEP	19	26%	15	8%	166	71%
Benchmark (ELA)	0	0%	184	99%	209	90%
Benchmark (Math)	0	0%	184	99%	207	89%
I/Leap ELA	67	91%	171	92%	199	85%
I/Leap Math	67	91%	172	92%	200	86%
TerraNova Math	15	20%	17	9%	0	0%
TerraNova Reading	15	20%	17	9%	0	0%
GRADE 6						
ELL	20	32%	19	11%	218	99%
FRPL	0	0%	17	10%	211	96%
Gender	33	52%	94	54%	119	54%
IEP	16	25%	13	7%	153	70%
Benchmark (ELA)	0	0%	170	98%	188	85%
Benchmark (Math)	0	0%	170	98%	193	88%
I/Leap ELA	53	84%	166	95%	179	81%
I/Leap Math	53	84%	168	97%	179	81%
TerraNova Math	22	35%	10	6%	0	0%
TerraNova Reading	22	35%	10	6%	0	0%

Michael & Susan Dell FOUNDATION

Firstline Schools (cont.)

Firstline	2010 N	2010 %	2011 N	2011 %	2012 N	2012 %
GRADE 7						
ELL	23	96%	26	18%	204	99%
FRPL	0	0%	24	16%	195	95%
Gender	14	58%	76	51%	113	55%
IEP	23	96%	18	12%	154	75%
Benchmark (ELA)	0	0%	148	100%	184	89%
Benchmark (Math)	0	0%	148	100%	187	91%
I/Leap ELA	24	100%	141	95%	171	83%
I/Leap Math	24	100%	141	95%	173	84%
TerraNova Math	0	0%	16	11%	0	0%
TerraNova Reading	0	0%	16	11%	0	0%
GRADE 8						
ELL	2	4%	22	100%	173	98%
FRPL	0	0%	17	77%	171	97%
Gender	28	62%	12	55%	98	55%
IEP	41	91%	20	91%	129	73%
Benchmark (ELA)	0	0%	21	95%	154	87%
Benchmark (Math)	0	0%	21	95%	157	89%
I/Leap ELA	41	91%	21	95%	145	82%
I/Leap Math	42	93%	21	95%	143	81%
TerraNova Math	0	0%	0	0%	0	0%
TerraNova Reading	0	0%	0	0%	0	0%

KIPP Empower Academy (KIPP LA)

KIPP LA	2010 N	2010 %	2011 N	2011 %	2012 N	2012 %
GRADE K						
ELL			105	91%	105	91%
FRPL			102	88%	102	88%
Gender			59	51%	59	51%
IEP			109	94%	109	94%
MAP Math			2	2%	116	100%
MAP ELA			3	3%	114	98%
GRADE 1						
ELL			106	92%	106	92%
FRPL			108	94%	108	94%
Gender			59	51%	59	51%
IEP			108	94%	108	94%
MAP Math			105	91%	112	97%
MAP ELA			105	91%	114	99%

Rocketship Education

	2010		2011		2012	
Rocketship	N	%	N	%	N	%
GRADE K						
ELL	103	53%	19	83%	484	73%
FRPL	171	89%	23	100%	543	82%
IEP	184	95%	21	91%	617	94%
Gender	103	53%	13	57%	344	52%
CST ELA	0	0%	0	0%	0	0%
CST Math	0	0%	0	0%	0	0%
MAP Math	192	99%	20	87%	622	94%
MAP ELA	192	99%	22	96%	621	94%
GRADE 1						
ELL	85	54%	17	57%	375	70%
FRPL	139	89%	25	83%	440	83%
IEP	153	97%	29	97%	488	92%
Gender	82	52%	18	60%	274	52%
CST ELA	0	0%	0	0%	0	0%
CST Math	0	0%	0	0%	0	0%
MAP Math	156	99%	29	97%	499	94%
MAP ELA	156	99%	30	100%	495	93%
GRADE 2						
ELL	76	57%	17	61%	313	65%
FRPL	121	91%	25	89%	399	83%
IEP	128	96%	27	96%	445	93%
Gender	79	59%	15	54%	242	50%
CST ELA	130	98%	27	96%	461	96%
CST Math	130	98%	27	96%	461	96%
MAP Math	132	99%	27	96%	454	94%
MAP ELA	132	99%	28	100%	452	94%
GRADE 3						
ELL	50	56%	11	61%	280	58%
FRPL	83	93%	17	94%	406	85%
IEP	83	93%	14	78%	440	92%
Gender	47	53%	10	56%	242	51%
CST ELA	88	99%	17	94%	458	96%
CST Math	88	99%	17	94%	454	95%
MAP Math	87	98%	18	100%	454	95%
MAP ELA	87	98%	18	100%	450	94%
GRADE 4						
ELL			3	75%	152	54%
FRPL			4	100%	245	88%
IEP			4	100%	250	89%
Gender			3	75%	151	54%

Rocketship Education (cont.)

Rocketship	2010 N	2010 %	2011 N	2011 %	2012 N	2012 %
GRADE 4 (CONT)						
CST ELA			4	100%	269	96%
CST Math			4	100%	269	96%
MAP Math			4	100%	268	96%
MAP ELA			4	100%	267	95%
GRADE 5						
ELL					55	55%
FRPL					90	90%
IEP					91	91%
Gender					52	52%
CST ELA					98	98%
CST Math					98	98%
MAP Math					98	98%
MAP ELA					98	98%

Summit Public Schools

Summit	2010 N	2010 %	2011 N	2011 %	2012 N	2012 %
GRADE 9						
ELL					351	81%
FRPL					285	66%
Gender					217	50%
IEP					181	42%
Benchmark (Alg)					385	89%
CST ELA					364	84%
CST Math					343	80%
G9 Placement					353	82%
MAP Math					288	67%
MAP ELA					329	76%

Michael & Susan Dell FOUNDATION

Statistical Modeling

The basic impact model compares the scores on a summative outcome measure between a treatment sample (where blended learning was used) and a comparison sample (where blended learning was not used), controlling for a measure of prior achievement. The prior achievement measure serves two functions: it increases the statistical power of the model by accounting for outcome variance (reducing the variance of the error term), and in some cases partially adjusts for pre-existing differences in the achievement distributions of the treatment and control groups.

Prior to fitting models, the two groups were compared on the basis of prior achievement measures to ascertain the similarity of the distributions of these measures. Following a guideline set out by the What Works Clearinghouse (What Works Clearinghouse. (2010). Procedures and standards handbook (Version2.1). Washington, DC: US Department of Education, Institute for Education Sciences.), if the distributions varied by more than .25 standard deviations, we were prepared to apply propensity score matching techniques to equate the two groups. However, we found no cases where the prior achievement measures varied by more than .25 standard deviations across the treatment and comparison samples.

MAIN IMPACT SPECIFICATION

A multiple linear regression model is used to model the expected outcomes by group, after controlling for prior achievement. When more than one school is present in a given group, the regression model is extended to a hierarchical linear model, indicating clustering within schools. The general impact model is specified as:

$$Y = ß_0 I_t + ß_1 ß_c + ß_2 I_t X + ß_3 I_c X$$

Where I_t and I_c are dichotomous indicator variables equal to 1 if a student belongs to the treatment or comparison condition, respectively, and X is a standardized (mean 0, variance 1) measure of prior achievement.

This model has no constant intercept term – rather, separate intercepts and slopes are fit for each treatment condition. Empirically, we have found that in some cases the slope parameters (b_2 and b_3) differ significantly between the two conditions, and thus a single slope parameter would result in a mis-specified model. Because the slopes of each regression line are potentially non-parallel, the estimated treatment impact could vary with values of X. We estimate the mean impact as the difference $b_0 - b_1$, which is the difference in expected outcome values (after adjusting for prior achievement) at the mean value of the prior achievement measure. Put more colloquially, $b_0 - b_1$ represents the impact for the average student.

SUBGROUP IMPACT SPECIFICATION

The impact-by-subgroup models specify a single common slope for the covariate term, along with a set of terms interacting the treatment indicator with indicators for each subgroup category. The general model is specified as follows:

$$Y = ß_0 I_{g1} I_t + ß_1 I_{g1} I_c + ß_2 I_{g2} I_t + ß_3 I_{g2} I_c + ß_4 X$$

Where I_{g1} and I_{g2} are dichotomous indicator variables equal to 1 if a student belongs to subgroup g_1 or subgroup g_2, respectively. In some cases more than two subgroups may be specified.

The four coefficients b_0 through b_0 represent the expected outcome values of cells in the 2 x 2 treatment-by-subgroup table when the standardized covariate is held constant at zero. We are mainly interested in the magnitude of the treatment contrast for each subgroup. Thus, we estimate values for the expressions $b_0 - b_1$ for the treatment contrast on group g_1, and $b_2 - b_3$ for the treatment contrast on group g_2, along with standard errors and test statistics for each contrast. A test for a statistically significant interaction is conducted on the value $(b_0 - b_1) - (b_2 - b_3)$; a significant non-zero value for this expression indicates a statistically significant interaction between the treatment and subgroup indicators.

Criteria for subgroup modeling. The candidate set of subgroups to model included gender, English language learner status, eligibility for free/reduced price lunch, and special education status. Subgroups were only modeled if the data were available and each subgroup represented more than 10% of the sample.

NWEA MAP VIRTUAL COMPARISON GROUP MODEL SPECIFICATION

Impacts analyses for two sites (KIPP Empower Academy and Rocketship Education) also included using the Virtual Comparison Group data provided by Northwest Evaluation Association (NWEA). Each student at each of these sites who took the NWEA MAP assessment in the fall and the spring in the target schools is matched with up to 50 similar students across the United States, based on similarity of fall test scores and school demographics. The mean of the spring scores of these similar students is considered a “virtual comparison” score.

We model the difference between the KIPP and Rocketship student scores and their virtual comparison scores using a simple impact model of the form:

$$D=\beta_0$$

When D is greater than zero, the KIPP students out-score their virtual comparison counterparts. For comparing subgroups, we expand the above model to:

$$D=\beta_0 I_{g1}+\beta_1 I_{g2}$$

This generates two difference estimates, one for each subgroup category.

SITE-SPECIFIC MODEL DETAILS

Below we list each site within the study, the outcomes analyzed, covariates used to control for pre-existing differences between groups, and the contrast(s) tested. Subgroups subjected to analysis are also listed.

SITE	OUTCOME(S)	GRADES	COVARIATE(S)	CONTRAST	SUBGROUP(S)
Alliance	Benchmark (ELA) CST (ELA, Math)	9,10 9	Grade 8 CST	**Treatment:** Tennebaum, Burton, Simon **Comparison:** School #5, Health Services, Ouchi	Gender, Median Prior Achievement
FirstLine	TerraNova (ELA, Math)	1, 2, 3	Prior Year’s TerrNova	**Treatment:** Arthur Ashe **Comparison:** John Dibert, S. J. Green, Langston Hughes Academy	Gender, Median Prior Achievement, IEP
KIPP Empower Academy	MAP VCG (ELA, Math)	K	—	**Treatment:** KIPP Empower Academy **Comparision:** Virtual Comparison Group	Gender, Median Prior Achievement
Rocketship Education	MAP VCG (ELA, Math)	K–5	—	**Treatment:** Rocketship Schools (5) **Comparision:** Virtual Comparison Group	Gender, Median Prior Achievement
Summit Public Schools	Benchmark (Math) CST (Math) NWEA MAP (Math)	9	Grade 9 Math Placement Exam	**Treatment:** Tahoma, Rainier vs. Everest **Comparision:** Tahoma, Rainier vs. Everest	Gender, Median Prior Achievement, ELL, FRPL

Model Parameter Reference Tables

In the tables that follow we list each impact model estimated, both for the main impact as well as impacts within subgroups. The table is organized by site. Within site we organize by type of model (main impact and specific subgroup), followed by the subject tested (ELA/Reading or Math) and the specific assessment and grade level.

The columns specify the coefficient of the fixed effects of the model, the standard error, and when appropriate the p-value. No p-value is specified for the coefficients of the group indicator variables or group-specific slopes; these are point estimates of group means, not differences (or any other statistic which would normally be tested against a null hypothesis value of zero).

In addition to the model coefficients, we also list the results of specific statistical tests (e.g., the contrast between the point estimates of two group means) as well as the number of cases used within the model.

KEY TO CONTRAST GROUP VARIABLE NAMES IN PARAMETER REFERENCE TABLE

Alliance College Ready-High Schools
0.blast_flag = Comparison Group
1.blast_flag = Treatment Group
FirstLine Schools
0.cond = Comparison Group
1.cond = Treatment Group
KIPP Empower Academy
Cons = Difference in RIT points between Treatment and Virtual Comparison Group
Rocketship Education
Cons = Difference in RIT points between Treatment and Virtual Comparison Group
Summit Public Schools
0.condition = Comparison Group
1.condition = Treatment Group

Michael & Susan Dell FOUNDATION

Alliance College-Ready High Schools

Alliance	B	SE	P
Main Effects			

ELA READING; BENCHMARK GRADE 9

Model Parameters			
0.blast_flag (Comparison Group)	3.20	0.05	
1.blast_flag (Treatment Group)	3.08	0.06	
0.blast_flag*std_prior_ela	0.43	0.03	
1.blast_flag*std_prior_ela	0.49	0.03	
schl_mean_prior_ela	-0.30	0.24	0.200
Counts			
N schls	6		
N Cases	718		
Statistical Tests			
Treatment vs Control	-0.12	0.08	0.136
Test of Differential Impacts by Prior Scores	0.06	0.04	0.150

ELA/READING: BENCHMARK GRADE 10

Model Parameters			
0.blast_flag	3.24	0.12	
1.blast_flag	3.11	0.15	
0.blast_flag*std_prior_ela	0.41	0.03	
1.blast_flag*std_prior_ela	0.52	0.03	
schl_mean_prior_ela	-0.04	0.39	0.916
Counts			
N schls	5		
N Cases	435		
Statistical Tests			
Treatment vs Control	-0.13	0.20	0.508
Test of Differential Impacts by Prior Scores	0.11	0.04	0.013

ELA/READING: CST GRADE 9

Model Parameters			
0.blast_flag	339.16	1.94	
1.blast_flag	325.08	2.17	
0.blast_flag*std_prior_ela	36.17	1.64	
1.blast_flag*std_prior_ela	39.60	1.64	
schl_mean_prior_ela	3.94	9.22	0.669
Counts			
N schls	6		
N Cases	721		
Statistical Tests			
Treatment vs Control	-14.07	3.01	
Test of Differential Impacts by Prior Scores	3.43	2.31	0.138

Alliance College-Ready High Schools (cont.)

Alliance	B	SE	P
MATH: CST GRADE 9			
Model Parameters			
0.blast_flag	339.44	13.18	
1.blast_flag	310.03	16.18	
0.blast_flag*std_prior_math	40.63	2.92	
1.blast_flag*std_prior_math	37.52	3.32	
schl_mean_prior_math	-42.66	38.32	0.266
Counts			
N schls	5		
N Cases	525		
Statistical Tests			
Treatment vs Control	-29.41	21.03	0.162
Test of Differential Impacts by Prior Scores	-3.11	4.42	0.482

Gender

	B	SE	P
ELA/READING: BENCHMARK GRADE 9			
Model Parameters			
0.blast_flag*0.gender	3.17	0.06	
0.blast_flag*1.gender	3.22	0.06	
1.blast_flag*0.gender	3.12	0.06	
1.blast_flag*1.gender	3.04	0.06	
std_prior_ela	0.46	0.02	
schl_mean_prior_ela	-0.31	0.23	0.186
Counts			
N schls	6		
N Cases	718		
Statistical Tests			
Treatment vs. Control: Female	-0.18	0.09	0.035
Treatment vs. Control: Male	-0.05	0.09	0.578
Treatment vs. Control: Female vs. Male (Interaction)	-0.13	0.08	0.093
ELA/READING: BENCHMARK GRADE 10			
Model Parameters			
0.blast_flag*0.gender	3.25	0.12	
0.blast_flag*1.gender	3.22	0.12	
1.blast_flag*0.gender	3.15	0.15	
1.blast_flag*1.gender	3.05	0.15	
std_prior_ela	0.46	0.02	
schl_mean_prior_ela	-0.07	0.38	0.847
Counts			
N schls	5		
N Cases	435		
Statistical Tests			
Treatment vs. Control: Female	-0.17	0.20	0.388
Treatment vs. Control: Male	-0.10	0.20	0.598

Alliance College-Ready High Schools (cont.)

Alliance	B	SE	P
Treatment vs. Control: Female vs. Male (Interaction)	-0.07	0.09	0.445

ELA/READING: CST GRADE 9

	B	SE	P
Model Parameters			
0.blast_flag*0.gender	338.92	2.65	
0.blast_flag*1.gender	339.58	2.55	
1.blast_flag*0.gender	324.55	2.94	
1.blast_flag*1.gender	325.84	2.89	
std_prior_ela	37.85	1.17	
schl_mean_prior_ela	4.29	9.91	0.665
Counts			
N schls	6		
N Cases	721		
Statistical Tests			
Treatment vs. Control: Female	-13.74	3.93	
Treatment vs. Control: Male	-14.37	4.05	
Treatment vs. Control: Female vs. Male (Interaction)	0.64	4.65	0.891

MATH: CST GRADE 9

	B	SE	P
Model Parameters			
std_prior_math	38.86	2.21	
schl_mean_prior_math	-42.20	37.45	0.260
0.blast_flag*0.gender	337.44	13.18	
0.blast_flag*1.gender	341.31	13.17	
1.blast_flag*0.gender	304.73	16.19	
1.blast_flag*1.gender	314.75	16.14	
Counts			
N schls	5		
N Cases	525		
Statistical Tests			
Treatment vs. Control: Female	-26.56	20.98	0.206
Treatment vs. Control: Male	-32.71	21.04	0.120
Treatment vs. Control: Female vs. Male (Interaction)	6.15	8.67	0.478

Median Prior

ELA/READING: BENCHMARK GRADE 9

	B	SE	P
Model Parameters			
std_prior_ela	0.40	0.03	
schl_mean_prior_ela	-0.29	0.23	0.205
0.blast_flag*0.median_prior_ela	3.14	0.06	
0.blast_flag*1.median_prior_ela	3.26	0.06	
1.blast_flag*0.median_prior_ela	2.99	0.07	
1.blast_flag*1.median_prior_ela	3.18	0.07	
Counts			
N schls	6		
N Cases	718		

Michael & Susan Dell FOUNDATION

Alliance College-Ready High Schools (cont.)

Alliance	B	SE	P
Statistical Tests			
Treatment vs. Control: Above Median	-0.08	0.09	0.389
Treatment vs. Control: Below Median	-0.15	0.09	0.081
Treatment vs. Control: Above vs. Below (Interaction)	0.08	0.08	0.338

ELA/READING: BENCHMARK GRADE 10

Model Parameters			
std_prior_ela	0.39	0.04	
schl_mean_prior_ela	-0.04	0.37	0.911
0.blast_flag*0.median_prior_ela	3.18	0.12	
0.blast_flag*1.median_prior_ela	3.29	0.12	
1.blast_flag*0.median_prior_ela	2.98	0.15	
1.blast_flag*1.median_prior_ela	3.23	**0.15**	
Counts			
N schls	5		
N Cases	435		
Statistical Tests			
Treatment vs. Control: Above Median	-0.06	0.19	0.741
Treatment vs. Control: Below Median	-0.19	0.19	0.318
Treatment vs. Control: Above vs. Below (Interaction)	0.13	0.09	0.135

ELA/READING: CST GRADE 9

Model Parameters			
std_prior_ela	31.36	1.95	
schl_mean_prior_ela	5.02	9.88	0.612
0.blast_flag*0.median_prior_ela	333.10	2.95	
0.blast_flag*1.median_prior_ela	345.33	2.98	
1.blast_flag*0.median_prior_ela	315.45	3.30	
1.blast_flag*1.median_prior_ela	335.76	3.42	
Counts			
N schls	6		
N Cases	721		
Statistical Tests			
Treatment vs. Control: Above Median	-9.57	4.01	0.017
Treatment vs. Control: Below Median	-17.66	3.95	
Treatment vs. Control: Above vs. Below (Interaction)	8.09	4.62	0.080

MATH: CST GRADE 9

Model Parameters			
std_prior_math	34.09	3.62	
schl_mean_prior_math	-43.83	38.85	0.259
0.blast_flag*0.median_prior_math	332.05	13.97	
0.blast_flag*1.median_prior_math	346.58	13.95	
1.blast_flag*0.median_prior_math	304.99	16.93	
1.blast_flag*1.median_prior_math	315.88	17.07	

Michael & Susan Dell FOUNDATION

Alliance College-Ready High Schools (cont.)

Alliance	B	SE	P
N schls	5		
N Cases	525		
Statistical Tests			
Treatment vs. Control: Above Median	-30.70	21.79	0.159
Treatment vs. Control: Below Median	-27.06	21.77	0.214
Treatment vs. Control: Above vs. Below (Interaction)	-3.64	8.85	0.681

Firstline Schools

Firstline	B	SE	P
Main Effects			

ELA/READING: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)

Model Parameters			
0.cond	599.45	1.60	
1.cond	602.14	2.96	
0.cond*std_prior_tn_reading	24.41	1.57	
1.cond*std_prior_tn_reading	20.47	3.42	
smean_std_prior_tn_reading	-17.03	6.31	0.007
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs Control	2.69	3.37	0.425
Test of Differential Impacts by Prior Scores	-3.94	3.74	0.292

MATH: BENCHMARK GRADE 4,5,6,7,8 (VS. COMPARISON)

Model Parameters			
0.cond	0.02	0.04	0.592
1.cond	-0.07	0.07	0.330
0.cond*std_prior_bench_math	0.82	0.03	
1.cond*std_prior_bench_math	0.93	0.04	
smean_std_prior_bench_math	-0.62	0.08	
Counts			
N schls	4		
N Cases	830		
Statistical Tests			
Treatment vs Control	-0.09	0.08	0.266
Test of Differential Impacts by Prior Scores	0.11	0.05	0.025

MATH: ILEAP GRADE 4,5,6,7,8 (VS. COMPARISON)

Model Parameters			
0.cond	-0.04	0.06	0.542
1.cond	-0.06	0.12	0.632
0.cond*std_prior_ileap_math	0.67	0.03	
1.cond*std_prior_ileap_math	0.67	0.08	
smean_std_prior_ileap_math	-0.97	0.14	
Counts			
N schls	4		
N Cases	756		
Statistical Tests			
Treatment vs Control	-0.02	0.14	0.893
Test of Differential Impacts by Prior Scores	0	0.08	0.968

MATH: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)

Model Parameters			

Michael & Susan Dell FOUNDATION

Firstline Schools (cont.)

Firstline	B	SE	P
0.cond	557.61	2.12	
1.cond	579.40	3.90	
0.cond*std_prior_tn_math	33.71	2.23	
1.cond*std_prior_tn_math	12.99	3.92	0.001
smean_std_prior_tn_math	-15.37	5.95	0.010
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs Control	21.78	4.44	
Test of Differential Impacts by Prior Scores	-20.72	4.43	

Gender

ELA/READING: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)

Model Parameters			
0.cond*0.gender_combined	597.35	2.17	
0.cond*1.gender_combined	602.04	2.41	
1.cond*0.gender_combined	603.12	4.06	
1.cond*1.gender_combined	601.40	4.32	
std_prior_tn_reading	23.41	1.47	
smean_std_prior_tn_reading	-16.78	6.31	0.008
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs. Control: Female	-0.64	4.94	0.897
Treatment vs. Control: Male	5.77	4.58	0.209
Treatment vs. Control: Male vs. Female (Interaction)	6.41	6.72	0.341

MATH: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)

Model Parameters			
std_prior_tn_math	28.99	2.02	
smean_std_prior_tn_math	-14.63	6.06	0.016
0.cond*0.gender_combined	558.25	2.90	
0.cond*1.gender_combined	556.80	3.22	
1.cond*0.gender_combined	582.77	5.44	
1.cond*1.gender_combined	574.42	5.82	
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs. Control: Female	17.62	6.64	0.008
Treatment vs. Control: Male	24.52	6.16	
Treatment vs. Control: Male vs. Female (Interaction)	6.90	9.06	0.446

IEP

Firstline Schools (cont.)

Firstline	B	SE	P
ELA/READING: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)			
Model Parameters			
0.cond*0.iep	602.89	1.69	
0.cond*1.iep	577.94	4.35	
1.cond*0.iep	604.52	3.23	
1.cond*1.iep	593.20	6.57	
std_prior_tn_reading	21.34	1.48	
smean_std_prior_tn_reading	-14.34	6.19	0.021
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs. Control: IEP	15.26	7.74	0.049
Treatment vs. Control: No IEP	1.63	3.64	0.654
Treatment vs. Control: No IEP vs. IEP (Interaction)	-13.63	8.55	0.111
MATH: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)			
Model Parameters			
std_prior_tn_math	27.63	2.11	
smean_std_prior_tn_math	-13.29	6.07	0.029
0.cond*0.iep	558.55	2.33	
0.cond*1.iep	551.70	6.00	
1.cond*0.iep	582.87	4.43	
1.cond*1.iep	562.57	8.99	
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs. Control: IEP	10.87	10.64	0.307
Treatment vs. Control: No IEP	24.32	4.99	
Treatment vs. Control: No IEP vs. IEP (Interaction)	13.45	11.74	0.252
Median Prior			
ELA/READING: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)			
Model Parameters			
std_prior_tn_reading	19.34	2.15	
smean_std_prior_tn_reading	-18.02	6.28	0.004
0.cond*0.med_std_prior_tn_reading	592.61	2.76	
0.cond*1.med_std_prior_tn_reading	606.68	2.86	
1.cond*0.med_std_prior_tn_reading	600.83	4.11	
1.cond*1.med_std_prior_tn_reading	603.63	4.71	
Counts			
N schls	4		
N Cases	602		
Statistical Tests			

Michael & Susan Dell FOUNDATION

Firstline Schools (cont.)

Firstline	B	SE	P
Treatment vs. Control: Above Median	-3.05	5.03	0.544
Treatment vs. Control: Below Median	8.22	4.49	0.067
Treatment vs. Control: Below vs. Above (Interaction)	11.27	6.74	0.095

MATH: TERRANOVA GRADE 1,2,3 (VS. COMPARISON)

Model Parameters			
std_prior_tn_math	27.44	3.00	
smean_std_prior_tn_math	-15.48	6.03	0.010
0.cond*0.med_std_prior_tn_math	552.23	3.93	
0.cond*1.med_std_prior_tn_math	562.48	3.67	
1.cond*0.med_std_prior_tn_math	586.59	5.42	
1.cond*1.med_std_prior_tn_math	567.90	6.76	
Counts			
N schls	4		
N Cases	602		
Statistical Tests			
Treatment vs. Control: Above Median	5.42	6.86	0.429
Treatment vs. Control: Below Median	34.35	6.03	
Treatment vs. Control: Below vs. Above (Interaction)	28.93	9.09	0.001

KIPP Empower Academy (KIPP LA)

KIPP Empower Academy	B	SE	P
Main Effects			
ELA/READING: MAP VCG GRADE K			
Model Parameters			
_cons	9.16	1.34	
Counts			
N Cases	83		
Statistical Tests			
MAP VCG (% >= VCG)	73.49		
MATH: MAP VCG GRADE K			
Model Parameters			
_cons	4.16	0.99	
Counts			
N Cases	85		
Statistical Tests			
MAP VCG (% >= VCG)	69.41		
Gender			
ELA/READING: MAP VCG GRADE K			
Model Parameters			
0.gender	5.59	2.01	0.005
1.gender	11.76	1.71	
Counts			
N Cases	83		
Statistical Tests			
MAP VCG, by Gender	6.17	2.64	0.019
MAP VCG (% >= VCG), Males	65.71		
MAP VCG (% >= VCG), Females	79.17		
MATH: MAP VCG GRADE K			
Model Parameters			
0.gender	6.42	1.46	
1.gender	2.42	1.28	0.059
Counts			
N Cases	85		
Statistical Tests			
MAP VCG, by Gender	-4	1.94	0.039
MAP VCG (% >= VCG), Males	83.78		
MAP VCG (% >= VCG), Females	58.33		
Median Prior			
ELA/READING: MAP VCG GRADE K			
Model Parameters			
0.median_prior_map_read	9.07	1.89	
1.median_prior_map_read	9.24	1.91	

KIPP Empower Academy (KIPP LA) (cont.)

KIPP Empower Academy	B	SE	P
Counts			
N Cases	83		
Statistical Tests			
MAP VCG, Interaction Lower Prior vs Higher Prior	-0.17	2.69	0.950
MAP VCG (% >= VCG), Lower Prior	78.57		
MAP VCG (% >= VCG), Higher Prior	68.29		

MATH: MAP VCG GRADE K

Model Parameters			
0.median_prior_map_math	4.92	1.45	0.001
1.median_prior_map_math	3.52	1.34	0.008
Counts			
N Cases	85		
Statistical Tests			
MAP VCG, Interaction Lower Prior vs Higher Prior	1.40	1.97	0.478
MAP VCG (% >= VCG), Lower Prior	69.23		
MAP VCG (% >= VCG), Higher Prior	69.57		

Michael & Susan Dell FOUNDATION

Rocketship Education

Rocketship	B	SE	P
Main Effects			

ELA/READING: MAP VCG GRADE K

Model Parameters			
_cons	4.74	0.90	
Counts			
N schls	5		
N Cases	574		
Statistical Tests			
MAP VCG (% >= VCG)	**68.64**		

ELA/READING: MAP VCG GRADE 1

Model Parameters			
_cons	2.98	1.98	0.132
Counts			
N schls	5		
N Cases	449		
Statistical Tests			
MAP VCG (% >= VCG)	62.58		

ELA/READING: MAP VCG GRADE 2

Model Parameters			
_cons	1.83	0.69	0.008
Counts			
N schls	5		
N Cases	397		
Statistical Tests			
MAP VCG (% >= VCG)	59.95		

ELA/READING: MAP VCG GRADE 3

Model Parameters			
_cons	3.57	1.31	0.007
Counts			
N schls	5		
N Cases	402		
Statistical Tests			
MAP VCG (% >= VCG)	65.17		

ELA/READING: MAP VCG GRADE 4

Model Parameters			
_cons	7.10	1.38	
Counts			
N schls	4		
N Cases	252		
Statistical Tests			
MAP VCG (% >= VCG)	79.37		

Rocketship Education (cont.)

Rocketship	B	SE	P
ELA/READING: MAP VCG GRADE 5			
Model Parameters			
_cons	8.23	1.11	
Counts			
N schls	2		
N Cases	90		
Statistical Tests			
MAP VCG (% >= VCG)	76.67		
MATH: MAP VCG GRADE K			
Model Parameters			
_cons	10.62	1.58	
Counts			
N schls	5		
N Cases	580		
Statistical Tests			
MAP VCG (% >= VCG)	85.52		
MATH: MAP VCG GRADE 1			
Model Parameters			
_cons	5.71	1.40	
Counts			
N schls	5		
N Cases	463		
Statistical Tests			
MAP VCG	5.71	1.40	
MAP VCG (% >= VCG)	70.41		
MATH: MAP VCG GRADE 2			
Model Parameters			
_cons	4.23	0.73	
Counts			
N schls	5		
N Cases	417		
Statistical Tests			
MAP VCG (% >= VCG)	73.14		
MATH: MAP VCG GRADE 3			
Model Parameters			
_cons	3.43	0.75	
Counts			
N schls	5		
N Cases	407		
Statistical Tests			
MAP VCG	3.43	0.75	
MAP VCG (% >= VCG)	68.30		

Rocketship Education (cont.)

Rocketship	B	SE	P
MATH: MAP VCG GRADE 4			
Model Parameters			
_cons	7.52	2.24	0.001
Counts			
N schls	4		
N Cases	261		
Statistical Tests			
MAP VCG (% >= VCG)	79.31		
MATH: MAP VCG GRADE 5			
Model Parameters			
_cons	12.20	1.43	
Counts			
N schls	2		
N Cases	90		
Statistical Tests			
MAP VCG (% >= VCG)	88.89		
Gender			
ELA/READING: MAP VCG GRADE K			
Model Parameters			
0.gender	4.69	1.00	
1.gender	4.80	1.01	
Counts			
N schls	5		
N Cases	574		
Statistical Tests			
MAP VCG, Interaction by Gender	0.11	0.87	0.896
MAP VCG (% >= VCG), Male	67.67		
MAP VCG (% >= VCG), Female	69.71		
ELA/READING: MAP VCG GRADE 1			
Model Parameters			
0.gender	2.40	2.02	0.233
1.gender	3.54	2.02	0.079
Counts			
N schls	5		
N Cases	449		
Statistical Tests			
MAP VCG, Interaction by Gender	1.14	0.98	0.244
MAP VCG (% >= VCG), Male	60.09		
MAP VCG (% >= VCG), Female	65.04		
ELA/READING: MAP VCG GRADE 2			
Model Parameters			

Rocketship Education (cont.)

Rocketship	B	SE	P
0.gender	2.01	0.85	0.018
1.gender	1.66	0.82	0.044
Counts			
N schls	5		
N Cases	397		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.35	0.97	0.721
MAP VCG (% >= VCG), Male	61.70		
MAP VCG (% >= VCG), Female	58.37		

ELA/READING: MAP VCG GRADE 3

Model Parameters			
0.gender	3.41	1.40	0.015
1.gender	3.72	1.39	0.008
Counts			
N schls	5		
N Cases	402		
Statistical Tests			
MAP VCG, Interaction by Gender	0.31	0.90	0.731
MAP VCG, Male	3.41	1.40	0.015
MAP VCG, Female	3.72	1.39	0.008
MAP VCG (% >= VCG), Male	62.05		
MAP VCG (% >= VCG), Female	68.12		

ELA/READING: MAP VCG GRADE 4

Model Parameters			
0.gender	8.77	1.45	
1.gender	5.16	1.49	0.001
Counts			
N schls	4		
N Cases	252		
Statistical Tests			
MAP VCG, Interaction by Gender	-3.61	1.23	0.003
MAP VCG (% >= VCG), Male	84.85		
MAP VCG (% >= VCG), Female	73.33		

ELA/READING: MAP VCG GRADE 5

Model Parameters			
0.gender	8.66	1.64	
1.gender	7.85	1.53	
Counts			
N schls	2		
N Cases	90		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.81	2.24	0.717
MAP VCG (% >= VCG), Male	78.57		
MAP VCG (% >= VCG), Female	75		

Michael & Susan Dell FOUNDATION

Rocketship Education (cont.)

Rocketship	B	SE	P
MATH: MAP VCG GRADE K			
Model Parameters			
0.gender	11.03	1.61	
1.gender	10.18	1.62	
Counts			
N schls	5		
N Cases	580		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.85	0.78	0.273
MAP VCG (% >= VCG), Male	85.67		
MAP VCG (% >= VCG), Female	85.36		
MATH: MAP VCG GRADE 1			
Model Parameters			
0.gender	5.83	1.47	
1.gender	5.59	1.47	
Counts			
N schls	5		
N Cases	463		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.24	0.87	0.780
MAP VCG (% >= VCG), Male	71.05		
MAP VCG (% >= VCG), Female	69.79		
MATH: MAP VCG GRADE 2			
Model Parameters			
0.gender	3.82	0.84	
1.gender	4.62	0.83	
Counts			
N schls	5		
N Cases	417		
Statistical Tests			
MAP VCG, Interaction by Gender	0.80	0.84	0.342
MAP VCG (% >= VCG), Male	69.65		
MAP VCG (% >= VCG), Female	76.39		
MATH: MAP VCG GRADE 3			
Model Parameters			
0.gender	3.80	0.87	
1.gender	3.06	0.86	
Counts			
N schls	5		
N Cases	407		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.74	0.84	0.380
MAP VCG (% >= VCG), Male	70.56		

Rocketship Education (cont.)

Rocketship	B	SE	P
MAP VCG (% >= VCG), Female	66.19		

MATH: MAP VCG GRADE 4

Model Parameters			
0.gender	8.72	2.27	
1.gender	6.05	2.30	0.009
Counts			
N schls	4		
N Cases	261		
Statistical Tests			
MAP VCG, Interaction by Gender	-2.67	1.15	0.020
MAP VCG (% >= VCG), Male	82.73		
MAP VCG (% >= VCG), Female	75.41		

MATH: MAP VCG GRADE 5

Model Parameters			
0.gender	12.26	1.79	
1.gender	12.14	1.71	
Counts			
N schls	2		
N Cases	90		
Statistical Tests			
MAP VCG, Interaction by Gender	-0.12	2.01	0.951
MAP VCG (% >= VCG), Male	90.48		
MAP VCG (% >= VCG), Female	87.50		

Median Prior

ELA/READING: MAP VCG GRADE K

Model Parameters			
0.median_prior_map_read	2.97	0.97	0.002
1.median_prior_map_read	6.47	0.97	
Counts			
N schls	5		
N Cases	574		
Statistical Tests			
MAP VCG (% >= VCG), Below Median	60.85		
MAP VCG (% >= VCG), Above Median	76.11		
MAP VCG, Interaction by Prior MAP Reading	3.50	0.86	

ELA/READING: MAP VCG GRADE 1

Model Parameters			
0.median_prior_map_read	3.47	2.03	0.087
1.median_prior_map_read	2.49	2.03	0.220
Counts			
N schls	5		
N Cases	449		

Rocketship Education (cont.)

Rocketship	B	SE	P
Statistical Tests			
MAP VCG (% >= VCG), Below Median	61.71		
MAP VCG (% >= VCG), Above Median	63.44		
MAP VCG, Interaction by Prior MAP Reading	-0.98	1.00	0.326

ELA/READING: MAP VCG GRADE 2

Model Parameters			
0.median_prior_map_read	1.89	0.83	0.022
1.median_prior_map_read	1.76	0.85	0.039
Counts			
N schls	5		
N Cases	397		
Statistical Tests			
MAP VCG (% >= VCG), Below Median	58.25		
MAP VCG (% >= VCG), Above Median	61.78		
MAP VCG, Interaction by Prior MAP Reading	-0.14	0.97	0.889

ELA/READING: MAP VCG GRADE 3

Model Parameters			
0.median_prior_map_read	3.54	1.39	0.011
1.median_prior_map_read	3.60	1.39	0.010
Counts			
N schls	5		
N Cases	402		
Statistical Tests			
MAP VCG (% >= VCG), Below Median	61.50		
MAP VCG (% >= VCG), Above Median	68.81		
MAP VCG, Interaction by Prior MAP Reading	0.06	0.91	0.944

ELA/READING: MAP VCG GRADE 4

Model Parameters			
0.median_prior_map_read	10.17	1.53	
1.median_prior_map_read	4.30	1.51	0.004
Counts			
N schls	4		
N Cases	252		
Statistical Tests			
MAP VCG (% >= VCG), Below Median	84.17		
MAP VCG (% >= VCG), Above Median	75		
MAP VCG, Interaction by Prior MAP Reading	-5.87	1.19	

ELA/READING: MAP VCG GRADE 5

Model Parameters			
0.median_prior_map_read	12.79	1.40	
1.median_prior_map_read	3.46	1.43	0.016
Counts			
N schls	2		

Rocketship Education (cont.)

Rocketship	B	SE	P
N Cases	90		
Statistical Tests			
MAP VCG (% >= VCG), Below Median	86.96		
MAP VCG (% >= VCG), Above Median	65.91		
MAP VCG, Interaction by Prior MAP Reading	-9.33	2.01	
MATH: MAP VCG GRADE K			
Model Parameters			
0.median_prior_map_math	9.27	1.60	
1.median_prior_map_math	12	1.60	
Counts			
N schls	5		
N Cases	580		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	2.73	0.77	
MAP VCG (% >= VCG), Below Median	77.85		
MAP VCG (% >= VCG), Above Median	93.13		
MATH: MAP VCG GRADE 1			
Model Parameters			
0.median_prior_map_math	6.75	1.48	
1.median_prior_map_math	4.59	1.49	0.002
Counts			
N schls	5		
N Cases	463		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	-2.15	0.90	0.016
MAP VCG (% >= VCG), Below Median	73.62		
MAP VCG (% >= VCG), Above Median	67.11		
MATH: MAP VCG GRADE 2			
Model Parameters			
0.median_prior_map_math	4.03	0.84	
1.median_prior_map_math	4.45	0.86	
Counts			
N schls	5		
N Cases	417		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	0.42	0.85	0.621
MAP VCG (% >= VCG), Below Median	68.35		
MAP VCG (% >= VCG), Above Median	78.39		
MATH: MAP VCG GRADE 3			
Model Parameters			
0.median_prior_map_math	3.59	0.89	
1.median_prior_map_math	3.26	0.88	
Counts			

Rocketship Education (cont.)

Rocketship	B	SE	P
N schls	5		
N Cases	407		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	-0.33	0.86	0.700
MAP VCG (% >= VCG), Below Median	64.32		
MAP VCG (% >= VCG), Above Median	72.12		

MATH: MAP VCG GRADE 4

Model Parameters			
0.median_prior_map_math	8.08	2.36	0.001
1.median_prior_map_math	6.87	2.38	0.004
Counts			
N schls	4		
N Cases	261		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	-1.21	1.16	0.295
MAP VCG (% >= VCG), Below Median	75.18		
MAP VCG (% >= VCG), Above Median	83.87		

MATH: MAP VCG GRADE 5

Model Parameters			
0.median_prior_map_math	15.12	1.64	
1.median_prior_map_math	9	1.68	
Counts			
N schls	2		
N Cases	90		
Statistical Tests			
MAP VCG, Interaction by Prior MAP Math	-6.12	1.89	0.001
MAP VCG (% >= VCG), Below Median	93.62		
MAP VCG (% >= VCG), Above Median	83.72		

Michael & Susan Dell FOUNDATION

Summit Public Schools

Summit	B	SE	P
Main Effects			

MATH: BENCHMARK GRADE 9

Model Parameters			
0.condition	65.48	1.52	
1.condition	64.46	1.50	
0.condition*std_g9_math_pct	13.20	1.41	
1.condition*std_g9_math_pct	19.90	1.63	
schl_mean_g9_math	54.92	11.53	
Counts			
N schls	4		
N Cases	307		
Statistical Tests			
Treatment vs Control	-1.02	2.36	0.667
Test of Differential Impacts by Prior Scores	6.71	2.16	0.002

MATH: CST GRADE 9

Model Parameters			
0.condition	362.95	4.50	
1.condition	402.19	4.53	
0.condition*std_g9_math_pct	50.46	4.23	
1.condition*std_g9_math_pct	85.13	5.18	
schl_mean_g9_math	102.62	33.85	0.002
Counts			
N schls	4		
N Cases	275		
Statistical Tests			
Treatment vs Control	39.24	7.07	
Test of Differential Impacts by Prior Scores	34.67	6.68	

MATH: MAP GRADE 9

Model Parameters			
0.condition	234.48	13.80	
1.condition	239.52	14.32	
0.condition*std_g9_math_pct	12.62	1.85	
1.condition*std_g9_math_pct	16.34	1.32	
schl_mean_g9_math	-16.83	234.94	0.943
Counts			
N schls	3		
N Cases	204		
Statistical Tests			
Treatment vs Control	5.04	5.41	0.352
Test of Differential Impacts by Prior Scores	3.72	2.27	0.101

ELL

Michael & Susan Dell FOUNDATION

Summit Public Schools (cont.)

Summit	B	SE	P
MATH: BENCHMARK GRADE 9			
Model Parameters			
std_g9_math_pct	15.22	1.11	
schl_mean_g9_math	54.66	11.55	
0.condition*0.ell	67.52	1.59	
0.condition*1.ell	55.92	3.60	
1.condition*0.ell	63.35	1.58	
1.condition*1.ell	63.72	3.13	
Counts			
N schls	4		
N Cases	307		
Statistical Tests			
Treatment vs. Control: ELL	7.80	4.82	0.105
Treatment vs. Control: Not ELL	-4.17	2.45	0.089
Treatment vs. Control: Not ELL vs ELL (Interaction)	-11.97	5.00	0.017
MATH: CST GRADE 9			
Model Parameters			
std_g9_math_pct	64.95	3.56	
schl_mean_g9_math	89.23	35.44	0.012
0.condition*0.ell	364.07	4.93	
0.condition*1.ell	373.25	11.23	
1.condition*0.ell	400.15	5.05	
1.condition*1.ell	386.97	9.99	
Counts			
N schls	4		
N Cases	275		
Statistical Tests			
Treatment vs. Control: ELL	13.72	15.22	0.368
Treatment vs. Control: Not ELL	36.08	7.70	
Treatment vs. Control: Not ELL vs ELL (Interaction)	22.36	15.82	0.157
MATH: MAP GRADE 9			
Model Parameters			
std_g9_math_pct	14.74	1.09	
schl_mean_g9_math	-10.43	234.73	0.965
0.condition*0.ell	236.65	13.80	
0.condition*1.ell	228.34	14.31	
1.condition*0.ell	239.97	14.32	
1.condition*1.ell	237.68	14.47	
Counts			
N schls	3		
N Cases	204		
Statistical Tests			
Treatment vs. Control: ELL	9.34	6.93	0.178

Summit Public Schools (cont.)

Summit	B	SE	P
Treatment vs. Control: Not ELL	3.32	5.46	0.543
Treatment vs. Control: Not ELL vs ELL (Interaction)	-6.02	5.26	0.252

FRPL

MATH: BENCHMARK GRADE 9

Model Parameters			
std_g9_math_pct	15.91	1.11	
schl_mean_g9_math	55.29	12.22	
0.condition*0.frl	66.35	1.63	
0.condition*1.frl	64.04	3.15	
1.condition*0.frl	64.04	1.92	
1.condition*1.frl	63.10	2.08	
Counts			
N schls	4		
N Cases	307		
Statistical Tests			
Treatment vs. Control: FRL	-0.94	4.00	0.814
Treatment vs. Control: Not FRL	-2.31	2.67	0.386
Treatment vs. Control: Not FRL vs FRL (Interaction)	-1.37	4.17	0.743

MATH: CST GRADE 9

Model Parameters			
std_g9_math_pct	63.84	3.48	
schl_mean_g9_math	86.75	36.36	0.017
0.condition*0.frl	364.21	4.98	
0.condition*1.frl	369.35	9.18	
1.condition*0.frl	405.23	5.95	
1.condition*1.frl	387.72	6.60	
Counts			
N schls	4		
N Cases	275		
Statistical Tests			
Treatment vs. Control: FRL	18.37	11.90	0.123
Treatment vs. Control: Not FRL	41.02	8.27	
Treatment vs. Control: Not FRL vs FRL (Interaction)	22.65	12.61	0.072

MATH: MAP GRADE 9

Model Parameters			
std_g9_math_pct	14.48	1.07	
schl_mean_g9_math	-10.20	230.61	0.965
0.condition*0.frl	235.45	13.55	
0.condition*1.frl	233.57	14.34	
1.condition*0.frl	242.61	14.09	
1.condition*1.frl	235.79	14.11	
Counts			
N schls	3		

Michael & Susan Dell FOUNDATION

Summit Public Schools (cont.)

Summit	B	SE	P
N Cases	204		
Statistical Tests			
Treatment vs. Control: FRL	2.22	7.17	0.757
Treatment vs. Control: Not FRL	7.16	5.40	0.185
Treatment vs. Control: Not FRL vs FRL (Interaction)	4.94	5.66	0.382

Gender

MATH: BENCHMARK GRADE 9

Model Parameters			
std_g9_math_pct	16.09	1.09	
schl_mean_g9_math	52.18	11.73	
0.condition*0.gender	65.04	2.00	
0.condition*1.gender	66.99	2.06	
1.condition*0.gender	63.05	2.04	
1.condition*1.gender	63.82	1.93	
Counts			
N schls	4		
N Cases	307		
Statistical Tests			
Treatment vs. Control: Female	-3.18	3.01	0.291
Treatment vs. Control: Male	-1.99	3.01	0.508
Treatment vs. Control: Male vs Female (Interaction)	1.18	3.74	0.752

MATH: CST GRADE 9

Model Parameters			
std_g9_math_pct	64.54	3.45	
schl_mean_g9_math	89.61	35.62	0.012
0.condition*0.gender	363.91	6.06	
0.condition*1.gender	366.81	6.32	
1.condition*0.gender	395.12	6.62	
1.condition*1.gender	399.72	5.98	
Counts			
N schls	4		
N Cases	275		
Statistical Tests			
Treatment vs. Control: Female	32.91	9.25	
Treatment vs. Control: Male	31.22	9.46	0.001
Treatment vs. Control: Male vs Female (Interaction)	-1.69	11.61	0.884

MATH: MAP GRADE 9

Model Parameters			
std_g9_math_pct	15	1.08	
schl_mean_g9_math	-10.18	234.49	0.965
0.condition*0.gender	232.34	13.85	
0.condition*1.gender	238.99	13.89	
1.condition*0.gender	240.48	14.35	

Summit Public Schools (cont.)

Summit	B	SE	P
1.condition*1.gender	238.93	14.32	
Counts			
N schls	3		
N Cases	204		
Statistical Tests			
Treatment vs. Control: Female	-0.06	5.76	0.992
Treatment vs. Control: Male	8.14	5.71	0.154
Treatment vs. Control: Male vs Female (Interaction)	8.20	3.95	0.038

Median Prior

MATH: BENCHMARK GRADE 9

Model Parameters			
std_g9_math_pct	17.78	1.99	
schl_mean_g9_math	54.17	11.40	
0.condition*0.median_prior	71.10	2.60	
0.condition*1.median_prior	60.48	2.47	
1.condition*0.median_prior	61.93	2.34	
1.condition*1.median_prior	66.78	2.38	
Counts			
N schls	4		
N Cases	307		
Statistical Tests			
Treatment vs. Control: Above Median	6.30	3.11	0.043
Treatment vs. Control: Below Median	-9.17	2.79	0.001
Treatment vs. Control: Below vs. Above (Interaction)	-15.47	3.68	

MATH: CST GRADE 9

Model Parameters			
std_g9_math_pct	64.70	6.10	
schl_mean_g9_math	98.94	33.79	0.003
0.condition*0.median_prior	377.17	7.73	
0.condition*1.median_prior	351.02	7.36	
1.condition*0.median_prior	384.69	6.92	
1.condition*1.median_prior	417.91	7.47	
Counts			
N schls	4		
N Cases	275		
Statistical Tests			
Treatment vs. Control: Above Median	66.89	9.48	
Treatment vs. Control: Below Median	7.51	8.45	0.374
Treatment vs. Control: Below vs. Above (Interaction)	-59.38	11.21	