Lessons from Five Years of Funding DIGITAL COURSEWARE

October 2014

SRI Education

SRI Education

SRI Education™ researchers address complex issues in education, learning, and human services. Multidisciplinary teams of education policy researchers, sociologists, psychologists, political scientists, statisticians, and others study education policy issues and develop research-based solutions to improve productivity and the quality of life at home and school and in the workplace.

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

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

Suggested citation:
Means, B., Peters, V., & Zheng, Y. (2014) *Lessons from Five Years of Funding Digital Courseware: Postsecondary Success Portfolio Review.* Menlo Park, CA: SRI Education.

Cover photo used with permission, courtesy of Montgomery College. Photo on page 4 used with permission, courtesy of Central Piedmont Community College, Chris Record.

SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025
Phone: 650.859.2000

www.sri.com/education

Lessons from Five Years of Funding DIGITAL COURSEWARE

Postsecondary Success Portfolio Review

Barbara Means
Vanessa Peters
Ying Zheng

October 2014

Contents

Introduction

With the long-term goal of dramatically improving students' degree completion rates, the Bill & Melinda Gates Foundation's Postsecondary Success initiative seeks to understand what is required for technology applications to produce positive student impacts at scale. The foundation believes that technology can play an important role in advancing its college completion agenda by making instruction more personalized to individual student needs, providing timely data to identify students requiring additional social and instructional supports, serving as the mechanism for spreading effective courseware and instructional practices across institutional and geographic boundaries, and making higher education more cost-efficient and affordable.[1]

From its beginning, the foundation's Postsecondary Success strategy targeted the problem of students entering college but not completing a degree or certificate program. A 2009 statement of strategy articulated the foundation's goal as doubling the number of students who earn a postsecondary credential with market value by age 26.[2] One of the three investment areas identified as a lever for addressing this goal was "new technology products and platforms that produce dramatic improvement in learning and completion rates and can be developed and adopted at scale." Subsequent foundation statements highlighted adaptive technologies that personalize a student's learning experience as the mechanism for achieving these outcomes: "One of our beliefs is that postsecondary experiences that better personalize learning for students have the potential to deliver better outcomes in terms of mastery and completion (with respect to actual completion rates, time to completion, and the cost of completion)."[3]

As the Postsecondary Success strategy enters its fifth year of learning technology investments, the foundation asked researchers at SRI Education to synthesize findings across the courseware projects it has funded to help identify lessons learned. This synthesis reports on types of courseware and related activities that have received Postsecondary Success funding, the characteristics of those courses and approaches that have and have not demonstrated positive outcomes for students, and the consistency of outcomes for those projects that have achieved positive impacts.

Postsecondary Success evaluation staff members selected 12 of their funded efforts for SRI's review (see Exhibit 1). These projects involved the development of courseware or other digital resources for individual courses rather than programs of study or entire degree programs. They also focused on courseware rather than new learning management systems, learning analytics, or early warning systems, although a few projects included elements of the latter as part of their activities.

---

[1] Bill & Melinda Gates Foundation, *Next Generation Learning*, 2010.

[2] Bill & Melinda Gates Foundation, *Postsecondary Success: Focusing on Completion,* 2009.

[3] Bill & Melinda Gates Foundation, *Request for Proposals: Adaptive Learning Market Accelerator Program,* 2013, p.1.

Exhibit 1. Postsecondary Success Technology Investments Reviewed in This Report

Technology Investment	Organization	Funding Date	Grant End Date	No. of Courses*	Planned to Measure Outcomes	Gates Funding	Abbrv. Name
NROC Developmental Math Redesign + DevMath Program to develop and distribute developmental math context and a personalized learning platform to provide instructional support	Monterey Institute for Technology in Education	Feb 2009	Dec 2014	1	Yes	$10.3M	NROC/ DevMath
Community College Open Learning Initiative Evaluated the use and effectiveness of CC-OLI environment for gatekeeper courses at 24 U.S. community colleges	Carnegie Mellon University	July 2009	Aug 2013	4	Yes	$2.5M	CC-OLI
Changing the Equation Whole-course redesign of developmental math sequences for 25 community colleges	National Center for Academic Transformation	Oct 2009	Nov 2013	36	Yes	$2.3M	NCAT/ CTE
Pathways Project: Quantway and Statway Courses Evaluated the effectiveness of an instructional system for developmental math and statistics courses	Carnegie Foundation for the Advancement of Teaching	June 2010	June 2014	2	Yes	$7.3M	Pathways
Next Generation Learning Challenges Wave I Multifaceted, collaborative initiative for leveraging technology to improve postsecondary completion rates for low-income college students (included 29 grantees)	EDUCAUSE	June 2010	Dec 2015	58	Yes	$17.9M	NGLC Wave I
DoL C3T Infrastructure + Open Course Library Partnership to provide technical, design, and implementation assistance to TAACCCT grantees	Creative Commons; WA-SBCTC	Apr 2011	Apr 2015	0	No	$12.8M	DoL C3T
OpenStax Developed a no-cost anatomy and physiology textbook that is peer-reviewed and edited by subject experts	Rice University	June 2011	June 2013	1	No	$0.8M	OpenStax
Planning and Implementation MITx/edX Explored efficacy of using a MOOC for credit in two community college computer programing courses	Massachusetts Institute of Technology	June 2012	March 2015	2	Yes	$1.1M	MITx/edX
Developmental and General Education MOOC 6 grants to assess feasibility of using MOOCs to provide free content within courses at accredited colleges	Various	Nov 2012	Nov 2013	9	Yes	$0.7M	Dev MOOC
MOOCs for Credit Research Evaluated educational potential of MOOCs as credit-bearing courses at degree or certificate-granting institutions	American Council on Education	Nov 2012	Apr 2014	10	No	$0.9M	ACE MOOC
University of Maryland MOOC Blended Course Project Tested various interactive learning platforms aimed at improving outcomes and reducing costs for students enrolled in traditional institutions	ITHAKA S+R	Nov 2012	July 2015	7	Yes	$1.8M	UMD Blended
Adaptive Learning Market Acceleration Program 17 grants for partnerships between postsecondary institutions and adaptive learning technology vendors to document and measure student learning outcomes	Various	June 2013	Feb 2016	7	Yes	$2.2M	ALMAP

**Number of distinct courses developed, reviewed, or evaluated by April 2014.*

The foundation selected projects for this review on the basis of their size and relevance to current trends in digital learning, including massive open online courses (MOOCs) and personalization. Three of the projects were sets of multiple grants or subgrants addressing a common goal. In total, the courseware projects reviewed by SRI encompassed 139 courses and represented approximately 90% of the foundation's investment in postsecondary courseware over the last five years.

It is important to keep in mind that this review reflects a window in time. Technology advances rapidly, and product features and approaches that are commonplace today were either just emerging or even unheard of in 2009, when the first of the grants reviewed here was awarded. To take a prominent example, MOOCs as we know them today did not really arrive on the scene until 2012, and the MOOCs that were the products of some of the Postsecondary Success grants reviewed here were using early versions of MOOC platforms that have since been revised.

The data sources for the meta-analysis in this report were the interim or final project reports submitted by grantees, interviews with project leaders, and research articles or data that projects provided in response to SRI's request.

The report is organized into the following sections:

1. **Project Descriptions** offer a synopsis of each project in the review.

2. **Synthesis Approach** describes the procedures and methods used for the analysis and synthesis of project reports.

3. **Characterizing the Portfolio** describes the features of the settings, technology and course design, implementation strategy, and measured student outcomes for the Postsecondary Success projects.

4. **Identifying What Works for Whom, How, and Why** presents data and findings from the courseware review as well as the design features and implementation strategies grantees adopted.

5. **Building Knowledge for the Field: Contributions and Challenges** discusses the Postsecondary Success courseware portfolio findings in the broader context of building knowledge for the field, including methodological issues that complicate data interpretation.

6. **Gaps in the Knowledge Base** highlights some of the important unanswered questions about how best to design and use digital courseware to improve outcomes for college students from low-income and underrepresented groups.

7. **Lessons Learned** summarizes some of the insights the foundation has gained into how to invest strategically in courseware projects and includes a conceptual framework for thinking about courseware investments and their evaluation.

8. **Conceptual Framework for Courseware Investments** considers the tensions between degree of innovation, implementation difficulty, scaling, and evidence of effectiveness and offers guidelines for determining appropriate evaluation approaches.

9. **Recommendations for Courseware Investments** articulates SRI's recommendations for philanthropic and governments organizations investing in digital courseware.

Project Descriptions

The 12 projects included in this review and the data they provided are described briefly below.

ALMAP

Under the Adaptive Learning Market Acceleration Program, 17 college campuses are using 10 adaptive learning technologies in course offerings in nine different subject areas in fully online, blended, and face-to-face implementation models. The terms of the ALMAP grant required that all participating campuses construct quasi-experimental treatment and control studies. However, several campuses that focus on serving older adult learners lacked data on baseline learning scores because they wanted to avoid alienating those learners. Outcomes measured include pre-post learning gains, final exam and course grades, course completion, and persistence to the next academic term. SRI was contracted to work with local evaluation teams to foster better evaluation practices, such as testing for baseline equivalence between groups and using instructors as their own controls. Across the seven ALMAP course implementations that have produced comparative outcome data so far, the ALMAP courses have produced slightly higher course completion rates. However, only one ALMAP course has had a statistically positive impact in terms of learning gains.

Changing the Equation (CTE)

Another large-scale, multicampus effort incorporating online learning was the course redesign work of the National Center for Academic Transformation (NCAT). This work was carried out with the explicit goal of reducing the costs of developmental (remedial) mathematics courses at community colleges while obtaining the same or better outcomes for students. NCAT promotes a particular model of instruction involving building the course around modularized online instructional software, holding class sessions in a computer lab or classroom, mastery learning with individualized pacing, and provision of on-demand assistance to students as they work with the software. The project's principal investigator, Carol Twigg, reported that all 32 CTE colleges that implemented the NCAT model as intended were able to improve student learning and reduce institutional costs.[4] (Cost savings came from increasing the size of sections and increasing the number of sections that full-time faculty were expected to teach.)

[4] Carol A. Twigg, *Improving Learning and Reducing Costs: Project Outcomes from Changing the Equation,* January 2012. http://www.thencat.org/Mathematics/CTE/CTE_Lessons.html

Community College Open Learning Initiative Project

The Open Learning Initiative (OLI) began in 2002 as an open educational resources R&D project at Carnegie Mellon University. The goal of the project was to develop a web-based learning environment where students who did not have access to a course instructor could achieve the same learning outcomes as students in traditional face-to-face courses. Successes in prior studies of OLI have shown that students using it achieved the same or higher learning outcomes in less time than peers who received traditional instruction.[5, 6] For the Community College Open Learning Initiative project, funded in 2009, OLI worked with faculty and software developers to develop introductory physiology, biology, psychology, and statistics courses for implementation in 24 community colleges. The OLI team used pre- and post-assessments and OLI system data to analyze achievement gains in OLI and non-OLI student groups. A regression model using propensity score matching indicated a positive effect for OLI students compared with a matched sample. This effect was not statistically significant, however, a finding that project leaders suggested might be due to the small matched sample size. Methodological challenges noted by project leaders include a large variation in how faculty prepared for class using OLI materials, how faculty graded student work, and how faculty required students to complete OLI activities. In addition, there were large differences in how faculty in OLI and non-OLI conditions implemented the pre- and post-tests, making it difficult to compare and judge the efficacy of the OLI materials.

Developmental and General Education MOOCs

Under this initiative, six institutions provided student outcome data from a MOOC using an existing technology platform (e.g., Udacity, Coursera) that provided open course access to a high volume of learners. Eligible courses were high-enrollment introductory-level courses that were designed with the same learning outcomes, content, and structure as a typical on-campus course at the awardee's institution. Eligible domain areas were developmental math, English language arts, social and behavioral sciences, physical and life sciences, and computer science. As a part of the award, all grantees were required to capture student and faculty activity data, anonymized student performance data, and anonymized student profile data for research purposes.

For some awardees, this was the first time a MOOC was implemented in their institution. A number of grantees developed course content from scratch (e.g., Cuyahoga Community College, University of Toronto), while others (e.g., Duke University, Wake Technical Community College) adopted existing course content or codeveloped course content with a MOOC vendor. Of the six grantees, five

[5] W. G. Bowen, M. M. Chingos, K. A. Lack, and T. I. Nygren, *Interactive Learning Online at Public Universities: Evidence from Randomized Trials,* May 22, 2012. www.sr.ithaka.org

[6] M. Lovett, O. Meyer, and C. Thille, "The Open Learning Initiative: Measuring the Effectiveness of the OLI Statistics Course in Accelerating Student Learning." *Journal of Interactive Media in Education,* May 2008. http://jime.open.ac.uk/jime/article/viewArticle/2008-14/351

provided descriptive measures such as course completion rate, student pass rate, student satisfaction, or the number of individuals who registered for the course or accessed course content. There were no experimental studies with comparative outcome data for any of the developmental and general education MOOC projects.

DoL C3T Infrastructure and Open Course Library

This project provides free support and services for institutions that are funded under the U.S. Department of Labor's Trade Adjustment Assistance Community College and Career Training (TAACCCT) grant program. DoL C3T offers assistance through the Open Professional Education Network (OPEN), which delivers a broad range of services to TAACCCT grantees through partnerships with Creative Commons, the Open Learning Initiative (OLI), the Center for Applied Special Technology, and the Washington State Board for Community & Technical Colleges. In their project interim report, DoL C3T grantees noted a number of challenges in providing this support. These included raising awareness of open education resources (OERs) and related practices, searching for and finding quality OERs, and the limited funds available to TAACCCT grantees for technical support. Preliminary results from TAACCCT grantees using the Open Course Library have not shown that use of OER results in significant changes in student success rates.

MOOCs for Credit Research

The American Council on Education (ACE) is engaged in research and implementation activities to evaluate and assess the landscape of MOOC delivery and make recommendations about the potential of MOOCs to contribute to degree completion in institutions of higher education. ACE is working with the University of Illinois Springfield's Center for Online Learning, Research, and Service to assess current practices in MOOCs, identify successful approaches, and create a potential pathway for prospective learners to apply open online learning toward credential and degree completion. At the time of this review, 10 MOOCs had been evaluated and had received ACE's recommendation for consideration of course credit by higher education institutions.

Next Generation Learning Challenges (NGLC) Wave I

The first NGLC grants were awarded in 2011 to 29 organizations with interventions designed to improve college students' course outcomes and prospects for degree completion by applying technology strategies. Twenty-three of these interventions involved changes to one or more courses, and 22 provided information on student outcomes for the course incorporating their intervention compared with prior versions of the same course. The core student outcome measure for most of the innovations was either course grade or "course success rate," defined as the proportion of enrolled students passing the course with a grade of C or better.

Overall, the NGLC innovations led to course completion rates on a par with those of instruction as usual. Obtaining equivalent outcomes was interpreted positively by those projects that were able to demonstrate cost savings, either for the institution offering the course or for the students, who were freed from the requirement to purchase a textbook because they could use open educational resources. NGLC results were somewhat more positive in terms of student learning outcomes such as score on an assessment or course grade. There was considerable variation across the projects in their impacts. Of the 17 NGLC projects that supplied learning outcome data that could be used to calculate an effect size, eight had statistically significant effects. Of those, seven favored the NGLC course and one favored the original version of the course prior to technology integration.

NROC Developmental Math Program

In 2009 the NROC Project received foundation funding to develop and distribute NROC Developmental Math, a multimedia-based remedial math program. (A second NROC grant, the value of which is included in Exhibit 1, funded development of the EdReady platform.) NROC Developmental Math was designed for use with students striving to meet college entrance requirements. Video, audio, adaptive practice problem sets, interactive simulations, an integrated textbook, and other instructional approaches were developed to support different learning styles and to engage students. NROC Developmental Math makes use of adaptive assessments to identify and address a student's proficiency gaps and can be installed to work with most learning management systems.

In three semesters in 2012-13, pilot studies were conducted to identify and document use cases for NROC Developmental Math and to measure student performance using the program. Student performance data was provided by 19 pilot sites, encompassing 16 secondary and postsecondary institutions in 11 states, with 31 instructors and 503 NROC-using students participating. In some pilot studies, classes using NROC Developmental Math were run in parallel with classes that did not use the program. In other cases, performance results using NROC Developmental Math were compared to historic measures of outcomes for the same class taught by the same instructor.

For this review, SRI included data only from pilot sites that had 25 students or more using NROC resources and that reported the percentage of students who passed with a C grade or higher for both the NROC students and a comparison group (either historic data or a parallel class). Of the 19 pilot sites, only 2 sites—both of which had historically high performing classes at 77% and 76% historic class pass rates, respectively—met these criteria.[7] While the two classes using NROC Developmental Math had improved class pass rates (85% and 79% pilot pass rates, respectively), impact estimates are based on only 102 NROC students and are therefore inconclusive. The impact of using the EdReady platform developed in the second NROC grant will be evaluated starting in 2015-16.

OpenStax Project

Based at Rice University, the OpenStax College project creates free textbooks that are developed and peer reviewed by educators and content experts to ensure they are readable, accurate, and meet the scope and sequence requirements of high-enrollment community college courses. The textbooks are available through Connexions, an open platform that enables instructors to customize the course texts for their needs. For this project, a total of 25 textbooks were developed in both online and PDF formats, as well as mobile and printed versions that will be available for a nominal fee. OpenStax College measures its success in terms of adoptions of textbooks and student savings estimated from what they would have paid for a textbook in a typical community or four-year college course. No cost savings data were provided for the Postsecondary Success courseware review.

Pathways Project

The Carnegie Foundation for the Advancement of Teaching's Pathways Project is a collaboration of more than 40 community and four-year colleges working on redesigned developmental mathematics courses. The Pathways Project addresses the problem of the low proportion of those students who need developmental mathematics who actually complete their developmental requirements and go on to earn credit in a college-level math course, a requirement for graduation. The project designed two different course sequences (or "pathways") representing alternative, intensified approaches. Both courses were built on the platform developed by the OLI and hence include extensive opportunities for online practice and assessment. In the Statway path, students encounter the basic mathematics they need as they are learning college-level statistics, and at the end of two semesters they have fulfilled their developmental math requirement and earned a college credit for statistics. In the Quantway pathway, students complete an accelerated developmental math experience in their first semester and then complete a credit-earning quantitative reasoning course the second semester. In the 2012–13 academic year, 1,439 students in 58 course sections started Statway. Before adopting the Pathways approach, the colleges implementing Statway saw only 6% of their entering students requiring developmental mathematics earn a college-level math credit within 12 months of continuous enrollment. Of students taking Statway in 2012–13, 68% successfully completed the first semester and 52% successfully completed both semesters, earning a college math credit. At the eight colleges using Quantway the same year, 51% of students successfully completed their developmental mathematics requirement in a single semester.[8]

[7] Results for all 19 pilot sites can be found at http://nrocm—ath.org/cms/wp-content/uploads/2010/10/Developmental-Math-report-7-17-13.pdf.

[8] Carnegie Foundation for the Advancement of Teaching, 2013.

University of Maryland Blended MOOC Project

In this project, ITHAKA S+R implemented and tested a variety of online learning platforms in blended MOOCs across 10 campuses within the University System of Maryland. The goal was to learn how emerging open technologies could be used to improve outcomes and reduce costs for students enrolled in traditional institutions. The project includes 12 side-by-side comparison tests and 10 case studies. Of the side-by-side comparisons, five were with Coursera, three with OLI, and four with Pearson. Because of the great variation in the MOOCs' design and implementation, in addition to the lead time required for testing, students were not randomly placed in different course formats. The range of disciplines in tested courses encompassed the humanities and social studies, math and statistics, communications, and computer science. Course sizes ranged from small seminars to large multisection introductory courses. For this review, side-by-side comparison data were provided for seven courses from four institutions. Overall, no significant differences were found in student learning outcomes in the hybrid courses. A small but statistically significant positive effect on likelihood of course completion was found for students in hybrid sections using OLI. In terms of satisfaction, students in hybrid sections gave lower course ratings and felt they learned less relative to students in traditional course sections.

Synthesis Approach

SRI developed a four-phase approach for extracting and synthesizing information from the Postsecondary Success courseware projects (see Exhibit 2). The goal of the analysis was to uncover key design features and implementation models among the reviewed projects and to relate them to student learning outcomes and scale of use.

Exhibit 2. Overview of Approach to PSS Portfolio Review

Development of a Portfolio Template

The first phase in the analysis was to create a template for gathering information across the 12 projects. To analyze the impact of technology investments on college completion and readiness more broadly, it was necessary to capture aspects of courseware design and implementation practices that might be associated with positive student outcomes. Recognizing that courseware products vary widely in purpose and functionality, we created a portfolio template to document such project characteristics as the technology affordance explored (e.g., game features), course subject area, and focus on helping low-achieving students. To develop a coding scheme, we drew on previous work describing categories and dimensions of online learning.[9] Exhibit 3 shows the features that were incorporated in the Postsecondary Success courseware portfolio analysis template, along with example values for each feature. The features were organized into three overarching categories: context of use, instructional design and technology features, and implementation practices.

Coding of Courseware Project Reports

Using project proposals and reports to the foundation supplemented by other published reports and information from project leader interviews, SRI coded each Postsecondary Success courseware project in terms of the features shown in Exhibit 3. In addition to coding for these descriptive features, we coded each project for reported outcome measures, including student outcomes such as course grade, course completion rate, and student satisfaction, as well as project or institutional outcomes such as scale achieved and cost savings for students.

[9] Barbara Means, Marianne Bakia, and Robert Murphy, *Learning Online: What Research Tells Us About Whether, When, and How.* New York, NY: Routledge, 2014.

Exhibit 3. Coding Categories Included in the Portfolio Analysis Template

Context of Use	Example Values
Field of use	Community college, four-year college, postsec training, self-initiated
Grantee organization type	Nonprofit higher ed, other nonprofit, consortium, gov't agency
Breadth	Course, portion of course, brief episode
Learners' preparation level	Weak, adequate, strong
Subject area	Math, statistics, science, social science & business, English, other
Instructional Design and Technology Features	**Example Values**
Pacing	Self-paced (open entry/open exit), class-paced, mixture
Dominant online pedagogy	Expository, practice environment, exploratory, collaborative
Source of feedback	Automated, teacher, peers
Peer engagement	Online discussion, help, joint product, peer grading, none
Online communication synchrony	Synchronous, asynchronous, both
System/platform type	MOOC, OLI, adaptive, LMS (learning management system), other
Instructor modifiability	Module selection, sequence, none
Individual learning paths	Mastery based, learner choice, none
Virtual experience	Virtual environment, avatars, teachable agent
Game features	Levels, badges, multiplayer game
Embedded online assessments	Multiple choice, extended response, delayed testing
Interoperability	Integrates with LMS, electronic grade book
Implementation Practices	**Example Values**
Modality	Fully online, blended with >50% online, blended with 25–50% online
Learning locations	Classroom, lab, home/dorm
Student: instructor ratio	≤ 35:1, 36–99:1, 100–299:1, 300–599:1, 600–999:1, ≥1,000:1
Online student: instructor interaction	High, medium, low
Online student: student interaction	High, medium, low

Interviews with Project Leaders

To fill in gaps in the portfolio analysis template data, we used information from interviews with project leaders as well as supplemental information or data the project leaders provided. In addition to findings related to student outcomes and achievement, the interviews produced important contextual information about the projects, including details about the student audience, the level of institutional support, and challenges or constraints that were specific to the local context of the project. A summary of each principal investigator interview appears in Appendix A.

Cross-Case Analysis and Outcome Data Meta-analysis

A quantitative meta-analysis was conducted using impact estimates from those projects that provided estimates of effect sizes or that reported quantitative data that enabled us to compute estimates of effect. The type and amount of data provided in project reports varied by grantee, as did detail on the specific models and methods the project used to estimate impacts. For project leader interviews, we performed a qualitative cross-case analysis of major themes in the areas of largest contribution, courseware design principles, implementation issues, and plans for sustainability.

Characterizing the Portfolio

Project coding using categories from the portfolio analysis template enabled us to characterize the context of use, instructional design and technology features, and implementation practices that were dominant across Postsecondary Success courseware efforts. Because many efforts included multiple courses, we coded each course or course-related activity for which a description was available, resulting in coded data for 139 different postsecondary courses or course implementations.

A majority of courses supported through Postsecondary Success funding were delivered in community colleges, and more than three-quarters of them focused on courses taken by lower achieving students (either developmental education courses or courses given in nonselective colleges). More than half the efforts (55%) involved designing or redesigning an entire course; others were less far-reaching, such as the creation of a supplemental instructional resource, an early alert system, or materials to support someone undertaking course redesign.

Much of the funded courseware was for mathematics or a similar subject like statistics; developmental mathematics was particularly common. Roughly 10% of the courses were MOOCs or hybrid courses incorporating MOOC resources. Exhibit B-1 in Appendix B shows the distribution of contextual features for the Postsecondary Success courses.

Exhibit B-2 in Appendix B shows the instructional design and technology features used in courseware funded by the Postsecondary Success initiative. ***The majority of courses had features consistent with the foundation's emphasis on personalization.*** More than 60% gave students flexibility by using an open entry/open exit approach to course timing. A majority also used individualized learning paths, in most cases through a mastery learning approach with online assessments used to determine whether the student was ready for new content. Almost all the funded courseware (90%) provided learners with automated feedback. Relatively few courses incorporated online synchronous communication, gaming features, or teachable virtual agents.

Regarding implementation practices, the Postsecondary Success courseware was used predominantly in a blended or hybrid mode without an active online instructor. ***Two-thirds of the courses engaged students by having them work on problems or answer questions as their primary online activity.*** Although the foundation funded a number of MOOC projects, the most common class size was in the 100 to 299 range. Few courses emphasized having students explore simulations or other interactive online resources, and few called explicitly for collaborating with peers. The distribution of codes for implementation practices is shown in Exhibit B-3 of Appendix B.

Finally, we coded the types of outcomes reported by the 12 Postsecondary Success courseware projects in this review. As shown in Exhibit 4, course completion (usually defined as finishing the course with a grade of C or better) was the most commonly reported outcome by far (83%). Other outcomes measured by roughly half of the projects were course examination scores, gains between pre- and post-tests, and student satisfaction. Very few projects examined longer term outcomes related to their interventions. Exceptions were several of the campuses working with NCAT, which measured student performance in the math course following their redesigned developmental math,[10] and the Pathways Project, which is looking at the number of college credits accumulated within two years of participating in a Pathways course.

Another perspective on the outcomes Postsecondary Success grantees reported is gained by considering the number of them that compared outcomes for different versions of their courseware or compared outcomes for their courseware with those for a different or conventional version of the course. Although qualitative statements about course improvement were common in project reports, ***only 7 of the 12 projects (58%) provided the kind of information needed to compute an estimate of their courseware's impact*** (i.e., an effect size) for any of their courseware implementations.

Exhibit 4. Outcomes Included In Postsecondary Success Project Reports

Project	Funding Date	Grant End Date	Course Completion	Pre-/ Post-Test	Exam Scores	Grades	Future Courses	Student Satisfaction	Socio-emotional	Cost Savings
NROC/ DevMath	Feb 2009	Dec 2014	✓	✓				TBD	TBD	
CC-OLI	July 2009	Aug 2013	✓	✓	✓					
NCAT/CTE	Oct 009	Nov 2013	✓	✓	✓		✓			✓
Pathways	June 2010	June 2014	✓	✓	✓	✓	✓		✓	
NGLC Wave I	June 2010	Dec 2015	✓	✓	✓	✓				
DoL C3T	April 2011	Apr 2015								
OpenStax	June 2010	June 2013	✓							TBD
MITx/edX	June 2012	March 2015	✓							
Dev MOOC	Nov 2012	Nov 2013	✓							TBD
ACE MOOC	Nov 2012	Apr 2014						✓		
UMD Blended	Nov 2012	July 2015	✓				✓	✓		✓
ALMAP	June 2013	Feb 2016	✓	✓			✓	✓		✓
% Reporting			83	50	33	17	33	25	8	25

TBD = To be determined (data were collected but not included in reports provided for this review).

[10] C. A. Twigg, http://www.thencat.org/Mathematics/CTE/CTE_Lessons.html

Identifying What Works for Whom, How, and Why

SRI used quantitative meta-analysis to summarize the available data on the effectiveness of the Postsecondary Success courseware projects in terms of student outcomes. Meta-analysis is a technique for synthesizing the results from a series of studies quantitatively. It has several advantages:

- It is more objective, systematic, and sophisticated than qualitative summaries or "vote-counting" of results from multiple studies because it provides a quantitative methodology for taking the strength of evidence from each empirical study into account;
- Meta-analysis produces synthesized effect estimates with considerably more statistical power than individual studies; and
- Differential effects related to different courseware features (moderators) such as a hybrid or fully online course design can be examined through follow-up tests for moderators of impact.

Meta-analysis requires creating a common metric—the effect size—that can then be averaged across studies and subsets of studies. An effect size is the difference between the average for the treatment (courseware) group and that for the comparison (business as usual) group divided by the standard deviation (a measure of how much individual scores differed from the average). Another way to think of effect size is as the impact of an intervention in standard deviation units. An effect size significantly larger than 0 indicates that the treatment group outperformed the comparison group. A significant negative effect size indicates that students in the comparison group performed better. If the treatment and comparison groups have identical performance on average, the effect size will be 0.

SRI analysts reviewed project reports and requested additional information from project leaders to obtain or estimate effect sizes for as many of the Postsecondary Success courseware investments as possible. Many projects included multiple courses or multiple implementations of the same course, and we used effect size data for each course and implementation when they were available. Following methodologists' recommendation not to combine effect sizes based on dichotomous variables (those with yes/no values like course completion) with those based on continuous variables (such as examination score or course grade),[11] we performed two separate meta-analyses: one using course completion rates and the other using outcome measures such as a grade or postassessment score. The overall results for these two analyses are shown in Exhibits 5 and 6.

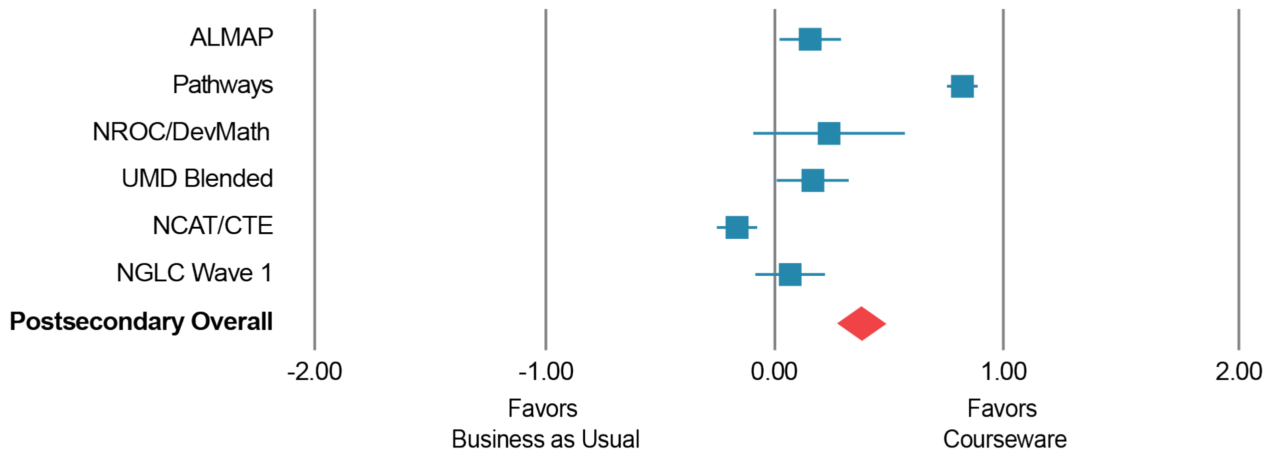
[11] Mark W. Lipsey and David B. Wilson, *Practical Meta-Analysis, Applied Social Research Methods,* Vol. 49. Thousand Oaks, CA: Sage, 2001.

Effectiveness of Postsecondary Success Courseware Projects

The meta-analysis on course completion used data from 94 course implementations. When examined at the level of the individual course implementation, there were 55 cases of no impact, 22 cases of a significant negative impact, and 17 cases of a significant positive impact. (Exhibits B-4 and B-5 in Appendix B provide effect estimates and confidence intervals at the level of individual courseware implementations for course completion and learning outcome measures, respectively. The details of the meta-analytic procedures are presented in Appendix C.)

We aggregated across individual course implementations to produce an average effect size for each of the seven projects as well as an overall average for the Postsecondary Success portfolio. The estimates of effect on course completion rate for four projects were close to 0, one project had a large positive effect, and one had a negative average effect (Exhibit 5). When the effects were aggregated at the portfolio level, the Postsecondary Success projects as a whole had a modest positive effect size of .37 on course completion rates. This effect size suggests that a student in a class using Postsecondary Success courseware is almost twice as likely as a student in a conventional version of the class to complete the course with a grade of C or better. However, this average effect estimate was influenced strongly by a single project with a large positive impact (Pathways). When we reran the course completion meta-analysis excluding the Pathways data, the effect estimate was very close to 0. Certainly the observed impacts on course completion have been more modest than the dramatic improvement in course and degree program completion rates envisioned in early articulations of the Postsecondary Success strategy. ***These findings suggest that there is more to learn about how to design and implement digital courseware in ways that produce positive impacts consistently across different settings.***

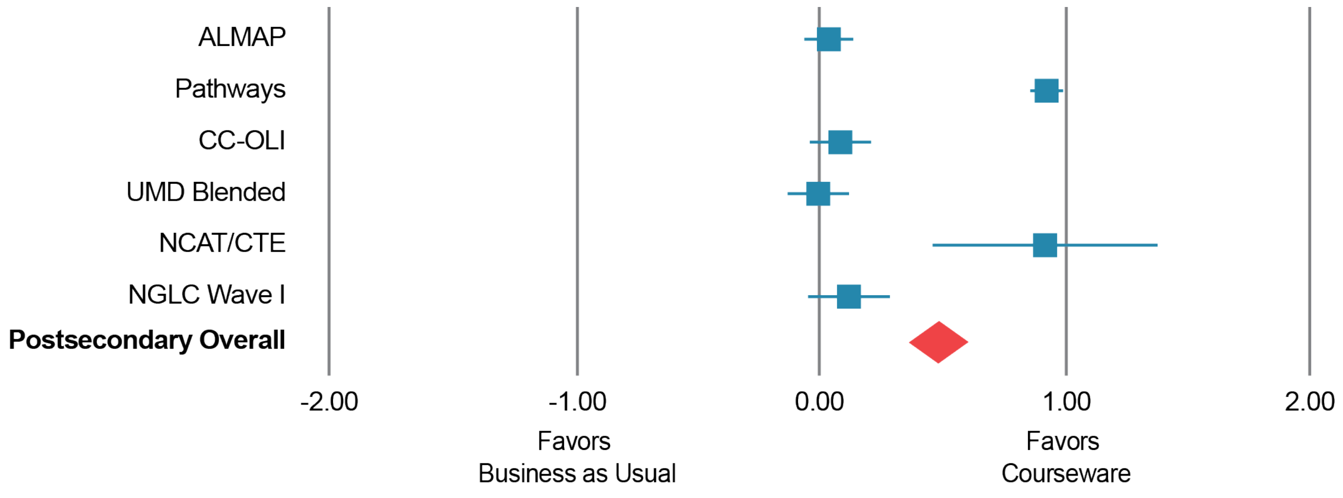
Exhibit 5. Impacts of Courseware on Course Completion Rate



In the meta-analysis on learning outcome measures available for 62 course implementations, we found a similar pattern of diverse but generally positive effects. ***The overall Postsecondary Success impact on learning outcomes was .47*** (see Exhibit 6). An impact of this size is equivalent to moving the average student in a class from a score of 50% to one of 68%. At the individual project level, the Changing the Equation and Pathways projects produced the greatest positive impacts, with effect estimates of .89 and .92, respectively. None of the other projects had an average effect estimate significantly different from 0. Although the NGLC Wave 1 project did not have a significant impact on average, its effect size was based on estimates from 13 different subgrants, several of which did have significantly positive impacts of moderate size.[12]

To obtain a sense of whether the positive effects shown in Exhibit 6 might be artifacts of poorly controlled quasi-experiments, we coded the available effect sizes for whether or not differences between students in the two conditions being compared had been statistically controlled.[13] This analysis showed that the average effect size for projects without statistical controls was .76, whereas that for projects with statistical controls for preexisting differences between student groups was .21, which although still significantly positive is considerably smaller in magnitude ($p < .001$). This finding suggests that ***project-reported comparative outcome data tend to overstate the courseware impact in cases where characteristics of students in the two types of courses are not measured and controlled.***

Exhibit 6. Courseware Impacts on Learning Outcome Measures



[12] California State University, Northridge and Missouri Community College Consortium.

[13] Of course, there may also be other potential confounding factors, such as differences between section instructors and course content.

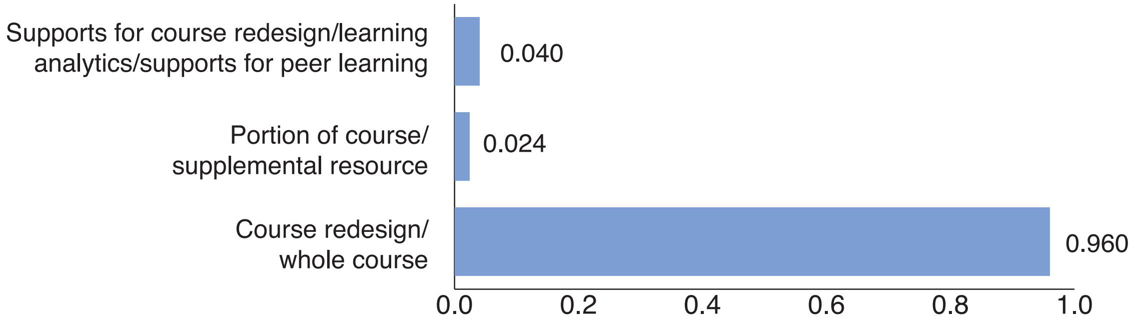
Moderators of Effectiveness

The Postsecondary Success initiative strives to produce insights that can shape future investments and uses of technology to enhance student success and higher education affordability. From this perspective, analyses of the use contexts, courseware design features, and implementation practices associated with greater effectiveness are of more interest than the average effect size. In previous work with the NGLC course outcome data, we found that we had more power to detect significant moderating variables when we used student learning outcomes measured by continuous variables (such as assessment score or course grade) than by course completion (which could be scored only as yes or no). Accordingly, we chose to use the 62 learning outcome effect estimates (shown in Exhibit B-5 in Appendix B) in our exploration of potential moderating variables.

We used the codes for the Postsecondary Success projects in our portfolio analysis template to test whether each variable was a significant moderator of course effectiveness. We tested only those variables for which we could construct logical contrasting groups (based on feature codes) with at least 10 cases in the smallest group; in many instances, different codes had to be combined into a logical group to achieve the required 10 cases.

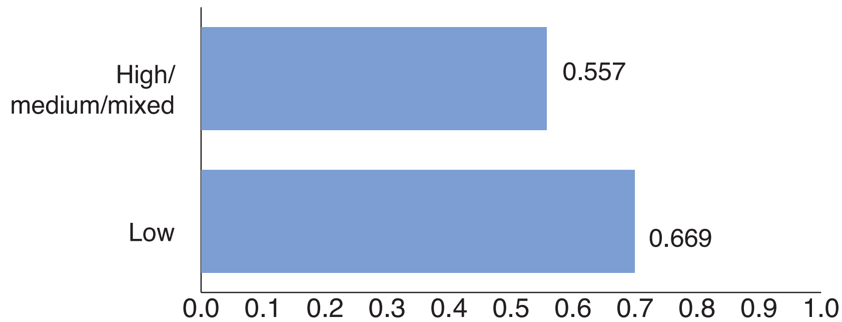
The use context codes for which the data met our criteria for moderator variable analysis were breadth, field of use, learners' preparation level, and subject area. We found that courseware learning outcomes varied with each of these dimensions. Implementation of whole course designs or redesigns produced significantly positive learning effects on average, whereas less intensive approaches (such as supplemental course resources or supports for the redesign process) did not (see Exhibit 7).

Exhibit 7. Effect Estimates by Courseware Role



Learning effects were larger in community colleges (ES = .93) than in four-year colleges (ES = .25). In marked contrast to several reports on analyses of other sets of online and blended courses,[14] we found that in the Postsecondary Success courseware portfolio, the estimated effect size for learning outcomes was as high for implementations involving students with weak prior academic preparation as for students with stronger records of academic achievement (Exhibit 8). The very large positive effects for Pathways and NCAT/CTE developmental mathematics courses appeared to drive this data pattern. Similarly, these projects contributed heavily to a significant subject area effect, with greater impacts for mathematics (ES = .82) than for other subjects (ES = .12).

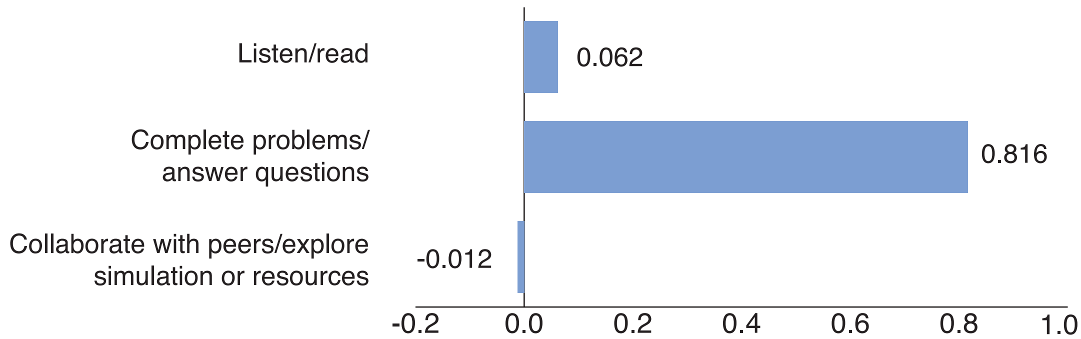
Exhibit 8. Courseware Effect Estimates by Learners' Preparation Level



The tested instructional design and technology features were dominant student role, pacing, dominant online pedagogy, individualized learning path, modality, and technology system/platform type. Courses in which the dominant role for students working online was solving problems or answering questions had larger positive effects than those where the dominant online activity was listening or reading (Exhibit 9).

[14] For example, see D. Xu and Shanna S. Jaggers, *Online and Hybrid Course Enrollment and Performance in Washington State Community and Technical Colleges,* CCRC Working Paper No. 31, New York: Columbia Teachers College, Community College Research Center, March 2011a; Xu and Jaggers, *The Effectiveness of Distance Education Across Virginia's Community Colleges: Evidence from Introductory College-level Math and English Courses, Educational Evaluation and Policy Analysis,* 33, 360, 2011b; and Figlio, Rush, and Yin, *Is It Live or Is It Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning,* NBER Working Paper 16089, Cambridge, MA: National Bureau of Economic Research, 2010.

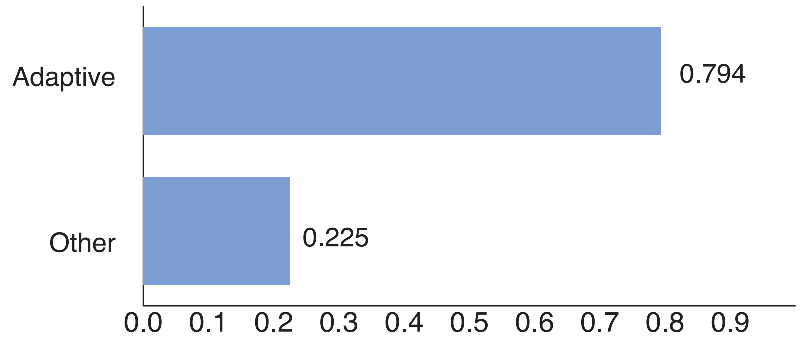
Exhibit 9. Courseware Effect Estimates by Dominate Student Role Online



Course implementations using individualized pacing had more positive impacts than those with class-based or a mixed form of pacing. In a related vein, courses using a mastery learning approach to individualizing students' learning paths had greater learning impacts than those that allowed learners to choose their own path through the material. In terms of online pedagogy, practice environments (which tended to be associated with mathematics courses) had more positive learning impacts than environments that called on students to collaborate or explore online resources.

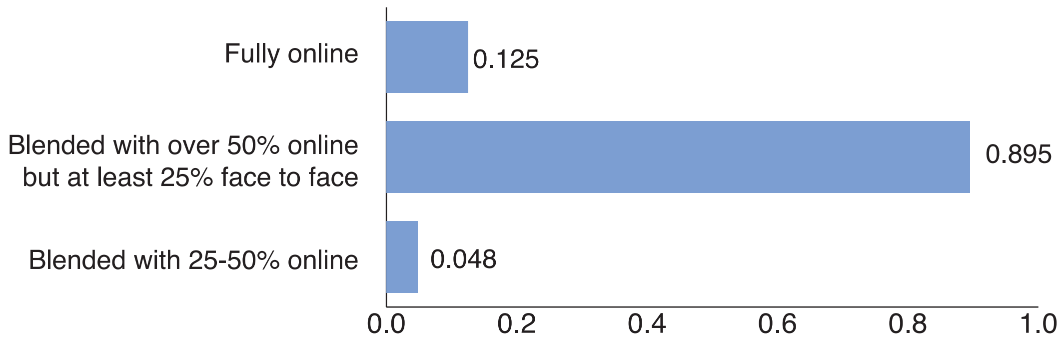
Too few MOOC implementations had learning outcome data for us to compare their outcomes with those of other technology system or platform types. The one conclusion the data did support is that ***adaptive learning technologies demonstrated larger learning effects than nonadaptive ones*** (Exhibit 10).

Exhibit 10. Courseware Effect Estimates by Technology Platform Type



Regarding implementation practices, we had enough effect estimates for contrasting groups on only the dimensions of modality and student-to-instructor ratio. Only 4 of the 62 learning effect estimates came from fully online implementations. Among the remaining effect estimates, blended courses in which an estimated 50% or more of a student's learning time was spent online had more positive impacts than blended courses with less time spent online (Exhibit 11).

Exhibit 11. Courseware Effect Estimates by Modality



In terms of student-to-instructor ratio, the relationship between course size and learning impacts appears nonlinear. Courses with 100 to 299 students had larger learning effect estimates than courses with either larger or smaller class sizes.

Exhibit 12 presents a summary of the significant modifier variables found in the Postsecondary Success courseware meta-analysis.

Courseware Design Features Cited by Project Leaders

In addition to identifying design features that were statistically significant moderators of courseware impacts, we asked the 14 project leaders we interviewed about the design principles that had guided their courseware design and those that they believed contributed to success. Design features endorsed in one or more interviews were:

Enabling students to set their own goals

As a project leader for the Duke University MOOCs explained,

> *Being able to pace yourself in the course was really important for us. In college, learning is top down, and the instructor hands out the materials. In a MOOC, the learners define for themselves what they want to do.*

Providing instructors with formative assessment data generated in the course of software use

Provision of detailed student progress data to students and instructors throughout the time the course was running was a feature of some recommended by the Pathways project leader. The software systems commonly used for developmental mathematics had this capability, as did some of the content-focused courses such as those of U-Pace and OLI. The MOOCs that had been implemented as of the time of this review, on the other hand, did not provide faculty with near real-time student learning data.

> *I almost think you should think about the faculty dashboard first and then build the student experience to provide that. The real power is the combined interaction of the student and teacher together. Faculty need a nimble way to understand what their students are struggling with so that they can adjust their instruction and provide targeted supports accordingly.*

Gaming features

Few Postsecondary Success courseware projects made use of badges, avatars, game mechanics, or online immersive experiences, but those that did regarded these technology features as important. As the Cuyahoga project leader explained,

> *The game mechanics were a key design feature; it was great for our audience. We called it the "X-MOOC." We're focused on community college students, and the game mechanics were part of the instructional design. This was done to motivate students to complete the course when they're not paying for it.*

Exhibit 12. Features Associated with More Positive Effects on Learning

1 Breadth	Effects were greater for projects either designing or redesigning an entire course than for those developing supplemental resources or early alert systems.
2 Field of use	Effect estimates were greater for projects implemented mainly in community colleges than in 4-year colleges.
3 Learners' preparation level	Effects were greater for projects targeting students with weak rather than moderate or advanced preparation.
4 Subject area	Mathematics courses had more positive effect estimates than courses in other subject areas.
5 Student: instructor ratio	Courses of medium enrollment size had more positive effects than the smallest and largest courses.
6 Pacing	Effects were larger for self-paced courses than for classes using cohort pacing or a mix of cohort and individualized pacing.
7 Dominant student role	Courseware in which the student's role was working on problems or answering questions had more positive effects than those where most time online was devoted to reading or listening to a video lecture.
8 Individualized	Courseware individualizing instruction on the basis of student performance on embedded assessments had more positive effects than those offering individualization based on student choice or no individualization.
9 Mastery based	Courseware determining when students are ready for new material by applying a standard of mastery had stronger learning effects than courseware allowing students to choose their own learning paths.
10 Adaptive technology	Learning systems that adapt to the individual learner had large learning impact estimates.
11 Modality	Effects tended to be more positive for courses using a blended learning model with more than half of the instruction occurring online.

Short duration

For MOOCs offered on an open basis rather than as part of a blended course, project leaders felt that keeping the duration short helped to preserve a higher level of participation. The Cuyahoga MOOC project leader, for example, reasoned,

> *The shorter duration of the MOOC course contributed to its success; it was six to eight weeks. In our MOOC, students can always go back in and start again. We noticed that attendance really starts to dive after four weeks, so we weren't interested in a 14-week course.*

Instructor moderation of online discussions

As noted above, most of the Postsecondary Success courseware designs called for little or no instructor activity online. However, one project explored the effects of increasing the level of the instructors' online involvement. Cuyahoga Community College compared two versions of its developmental math MOOC with different degrees of instructor involvement and interactivity. In the first version of the MOOC, the discussion forums were not moderated, but in the second, the discussion forums were moderated and instructors sent out emails to connect with students who had not been participating. Students enrolled in the moderated version of the developmental math MOOC were more successful in terms of engagement and course persistence than students enrolled in the nonmoderated version. This finding is consistent with other online learning research emphasizing the importance of "instructor presence."[15]

Face-to-face classroom interactions to supplement the online learning

Carol Twigg, leader of the NCAT/CTE project, for example, argues that a blended approach with face-to-face access to an instructor or teaching assistant is necessary for low-achieving students:

> *We believe the face-to-face component is critical. We don't think teaching developmental math in a fully online environment will work for most students. Human contact, keeping them on point and encouraging them, is critical.*

[15] D. R. Garrison and J. B. Arbaugh, Researching the Community of Inquiry Framework: Review, Issues, and Future Directions. The Internet and Higher Education, 10(3), 157–172, 2007, doi:10.1016/j.iheduc.2007.04.001.

Building Knowledge for the Field: Contributions and Challenges

This section of the report provides a broader context for the major findings from our review by discussing some of the major challenges in the field of evidence-based technology-supported learning.

Contexts and Student Types

Given its long-term goal of dramatically increasing the number of young adults earning a postsecondary credential or degree, Postsecondary Success has a focus on innovations that affect those students least likely to finish college—low-income students, African American and Hispanic students, and first-generation college goers. These students are more likely to begin higher education at a community or technical college than at a four-year college.

The foundation is most interested in technology-supported courseware innovations that are effective with these students, but there is some controversy in the literature about the use of online learning with low-achieving and community college students who have the option to take courses in a conventional classroom-based format. Analyses of institutional data maintained by community college systems in Washington and Virginia led Xu and Jaggers of Columbia's Community College Research Center (CCRC) to conclude that low-income, less well-prepared students are likely to suffer when they take courses online.[16] They found, for example, that Virginia Community College students taking their first math course online were less likely to complete it successfully than those taking it in a classroom (67% compared with 73%).

In contrasting CCRC's negative findings with the positive findings of the NCAT/CTE and Pathways developmental mathematics interventions described in this report, it is important to consider the characteristics of the two sets of courses. The CCRC research team conducted an analysis of a sample of the online courses offered by one of the community college systems for which they had previously analyzed student outcome data. They found that the online courses these students had taken in 2004 were generally of poor quality and a far cry from what is available today.[17] Some of the so-called online courses consisted solely of posting a course syllabus online and collecting assignments through the course management system. Many of the courses were primarily textbook driven and contained no multimedia elements.

[16] D. Xu and S. S. Jaggers, *Online and Hybrid Course Enrollment and Performance in Washington State,* 2011a; and E*ffectiveness of Distance Education Across Virginia's Community Colleges,* 2011b.

[17] R. H. Bork and Z. Rucks-Ahidiana, *Virtual Courses and Tangible Expectations: An Analysis of Student and Instructor Opinions of Online Courses.* Paper presented at the Annual meeting of the American Educational Research Association, 2012.

It is also important to remember that the data Xu and Jaggers analyzed contrasted fully online courses with classroom-based courses. One might expect quite different results for low-income, minority, and low-achieving students when blended courses are contrasted with courses without online elements. The data from the NCAT/CTE project and from the Carnegie Foundation's Pathways Project suggest that blended courses with significant online learning components can be effective with community college students placed in remedial mathematics. Similarly, the NGLC Wave 1 and ALMAP projects reporting separate impact estimates for low-income students almost always found positive effects for this subgroup if they had positive effects overall.

The Postsecondary Success courseware projects have demonstrated that the incorporation of both mastery learning components and interactive face-to-face instruction in a blended learning model is associated with improved student learning outcomes in mathematics. The Postsecondary Success initiative has funded the redesign of developmental and gateway mathematics courses to incorporate technology on many campuses. Estimates of impact have come from 22 NGLC Wave I projects, 5 ALMAP projects, 18 CTE projects, 1 NROC project, 4 Statway campuses, and 3 Quantway campuses. Across these projects, the majority have reported positive impacts in terms of one or more student outcomes (typically a mathematics assessment, course grade, or course completion). Nearly all of these have involved use of mastery learning principles to individualize the pacing of student learning for at least some portion of the redesigned course. Almost all have involved some degree of face-to-face engagement as well as use of instructional software.

The Challenge of Obtaining a Fair Outcome Measure

The findings described above add considerably to what was available in the literature circa 2008. Nevertheless, we recognize the limitations of many of these studies. For a variety of reasons, very few of the individual NGLC Wave 1, ALMAP, and NCAT/CTE projects set up random-assignment experiments to provide a rigorous test of the effectiveness of their online learning innovations.

In some cases, technology-based and conventional sections of a course ran concurrently so that student outcomes from the same academic term could be compared. More often, the best data available were course outcomes for the newly designed course compared with those for the same course taught in prior years to different student cohorts (and not necessarily by the same instructor). Hence, the individual projects' reported outcomes could easily be confounded by differences between the students taking the two versions of the course or by differences in instructors that have nothing to do with the online or blended learning interventions themselves. The ALMAP projects provided student data on demographic characteristics and prior achievement of the students involved in the quasi-experiment, but the NGLC Wave 1 and NCAT/CTE projects typically did not collect such data.

Another challenge for the accumulation of research findings with respect to the effectiveness of different approaches to online and blended learning is the difficulty identifying outcome measures that can be used across different conditions to draw a fair comparison. Course grades or course completion with a C or better (a categorical variable based on course grade) are typically the only measures available in cases where the implementation of the redesigned course is not part of a separately funded evaluation study. Grade-based outcomes are clearly vulnerable to inconsistencies in instructor grading criteria and practices. Differences in grading stringency between instructors whose course sections are being compared should even out between treatment and control conditions when data from very large numbers of implementations are examined but can certainly lead to invalid inferences about the impact of any small number of course implementations.

Another common problem is differential retention in the conditions being compared. The very high rate of nonparticipation among those who sign up for MOOCs is the extreme example, but in online courses generally, the dropout rate is significantly higher than in traditional college courses.[18] To take an extreme example, if 80% of the students in an online section of a course decide to discontinue taking the course, then comparing the performance of the 20% who stick it out with that of students in a traditional section of the course where only 10 to 15% drop out leads to biased estimates of relative effectiveness.

Jaggers argues that because of differential course attrition rates for online and face-to-face courses, straight-up comparisons of end-of-course learning assessments are biased in favor of online courses.[19] As lower achieving students are more likely to drop a course, large discrepancies in course completion rates produce biased samples by the course's end. The problem with attempts to control for such differences statistically is that analysts cannot be sure they have not over- or under-corrected.

NCAT's Twigg, on the other hand, argues that straight-up comparisons of mastery-based blended learning courses and traditional courses is unfair in ways that are disadvantageous to the former.[20] Use of mastery learning means that students do not progress to the next course module until they have really mastered the current one, as demonstrated by a score at a predetermined high level (typically 75–85%) on the current module's assessment. In contrast, a student earning a 70% on each module in a traditional course will earn a course credit with a C. NCAT has found

[18] S. S. Jaggers, *Online Learning: Does It Help Low-Income and Underprepared Students?* CCRC Brief, No. 52, New York: Columbia Teachers College, Community College Research Center, March 2011

[19] Shanna S. Jaggers, *Online Learning.* CCRC Brief, No. 52, New York: Columbia Teachers College, March 2011.

[20] Carol A. Twigg. *Improving Learning and Reducing Costs: Program Outcomes from Changing the Equation,* The National Center for Academic Transformation.

that many developmental math students at community colleges need more than a single academic term to master all they have to learn, and hence they are likely not to have earned the credit for a mastery-based course by the end of term. She argues instead for using a measure of making progress rather than course completion for comparing redesigned and traditional courses. (Making progress is intended to be equivalent to course completion with a C and was defined variously by different NCAT/CTE campuses with such criteria as "completed 70% of modules at 85% mastery" or "completed 75% of modules at 80% mastery.") When the Changing the Equation courses were compared with prior versions of the same courses in terms of students' rates of passing with a C or better, 10% of them had significantly better pass rates and 34% had significantly lower pass rates. When the making progress metric was used as the outcome, 42% of the redesigned courses had significantly better outcomes and only 12% had significantly poorer outcomes. Clearly, the outcome measure makes a difference. When comparing course completion data, it is important to know whether one or both of the courses being compared has a mastery learning design. Also important is having data on the proportions of the students who started the traditional and the redesigned courses who were still participating and available to take whatever assessment was used as a learning outcome.

An even more complex issue related to identifying appropriate outcomes is that of near- versus long-term outcomes. In general, those who have learned material to a greater extent will be more likely than those with lower degrees of mastery to recall and use the material in later learning. Yet there is a body of research showing that some direct-instruction approaches that produce more efficient initial learning result in less long-term retention and subsequent use of that learning than approaches involving guided discovery.[21] In the absence of longer term measures of performance in subsequent courses, we cannot say whether or not the course design features and implementation practices that are associated with stronger near-term outcomes are really best for students' longer term persistence and future learning.

Thus, analytic choices, in addition to characteristics of course design, students, and organizational practices, influence observed student outcomes. Postsecondary Success projects conducting their own evaluations of their innovation's effectiveness tended to report the outcome measure their product looked best on without providing information on other types of outcomes.

This challenge of reporting outcomes was acknowledged in a number of interviews with project leaders. One MOOC grantee, for example, described the difficulty of reporting findings when there was no consensus about what is considered a meaningful outcome. As one interviewee explained,

> *MOOCs are a useful tool, but the fact that 30,000 signed up and 300 finished is irrelevant. What matters is the output. Was our MOOC working with a specific population with special needs? We don't know. And in terms of efficacy, we weren't set up in a way that we could measure it.*

---

[21] A. F. Wise and K. O'Neill, Beyond More Versus Less. In S. Tobias and T. M. Duffy (eds.), *Constructivist Instruction: Success or Failure?* New York: Routledge, 2009.

We conclude that it is inefficient and ineffective for a funding agency to rely solely on grantee choices about what outcome data to report and on grantee efforts to collect these data (often post hoc) from higher education institutions using their courseware.

The Challenge of Distilling Generalizable Principles

Clearly, there are many different kinds of courseware and online and blended learning implementation models, and they should not be lumped together indiscriminately for labeling as either "effective" or "ineffective." But the status quo is doing little to provide an accumulation of evidence about the instructional design principles that enhance different kinds of learning for different kinds of learners. Part of this deficiency stems from the sheer complexity involved in designing a piece of courseware, a complexity that then is multiplied once one considers all the different possible approaches for incorporating that software into a course. The Pittsburgh Science of Learning Center at Carnegie Mellon University has attempted to take on this problem by elucidating a set of instructional principles based on the learning theory underlying its tutoring systems and testing those principles in the context of cognitive tutors and OLI courses in different subject areas. Exhibit 13 shows the kinds of insights that can be gained from a systematic effort to harvest findings from the student outcome data generated when learning software is designed deliberately using learning science principles.

The matrix of design features of Postsecondary Success courseware presented earlier in Exhibit 3 is an attempt to apply a similar strategy, albeit at a coarser grain size. Our efforts to identify design features for the Postsecondary Success courseware were hindered by the fact that most college courses are developed without involving learning researchers, and the developers' design choices often go undocumented and untested. This state of affairs exists not because course developers are unconcerned with the quality of their courseware, but simply because their expertise and research interests usually lie in fields other than learning research, and they are more concerned with producing courseware they can use with their classes than with building knowledge about blended and online learning.

There were a few examples of Postsecondary Success courseware projects exploring specific design principles, as described in the previous section, but these principles were tested in the context of just a single course rather than across multiple courses or student populations. Design principles offer an alternative way of focusing learning technology investments.

Exhibit 13. Pittsburgh Science of Learning Center Instructional Principles

Principle	Principle Application	Subject Matter								
		CV	FA	EA	AE	GR	CR	HS	PP	PC
Optimized scheduling	Selection of practice instances based on prior statistics and on each student's experience with each target knowledge component.	+				+				
Timely feedback	Providing an evaluative response (e.g., correct or incorrect) soon after a student's attempt at task or step.		+		+			+		
Feature focusing	Guiding the learner's attention (focus) on valid or relevant features of target knowledge components.	+	+							
Worked examples	Worked examples are interleaved with problem solving practice (as opposed to practice that is all problem solving).				+	+	+			?
Prompted self-explanation	Encouraging students to explain to themselves parts of instruction (steps in worked example or sentences in a text).			0	+	+			+	?
Accountable talk	Encouraging classroom talk that is accountable to accurate knowledge, rigorous reasoning, and the classroom community by using six talk moves (question and response patterns).									+

Note: CV = Chinese vocabulary; FA = French articles; EA = English articles; AE = algebra equations; GR = geometry rules; CR = chemistry rules; HS = help-seeking skills; PP = physics principles; PC = pressure concept

+ = Principle found to enhance learning of this type of knowledge component.

0 = Principle tested for this knowledge component and no effect found.

? = Principle tested with this knowledge component with inconclusive results.

Source: B. Means, M. Bakia, and R. Murphy, *Learning Online*. New York: Routledge, 2014.

Gaps in the Knowledge Base

Educators, policymakers, and funders want information on the effectiveness of particular interventions incorporating online learning for particular purposes, types of students, and contexts. Whereas it is now well established that online courses are effective for some purposes and types of learners,[22] we are still some distance from being able to answer the kinds of more specific questions that college administrators have, such as

1. How should digital courseware be designed to make it reusable in different colleges and with different student populations?

2. Does online courseware put community college students who are English learners at a disadvantage?

3. Do courses incorporating online mastery learning modules result in more or less deep conceptual learning than conventional versions of the same course?

4. How can instructors use some of the data captured by learning systems to improve their teaching?

Answering these more specific questions about the purposes, student populations, and contexts in which specific interventions involving digital courseware can contribute to improved outcomes would help college leaders and faculty make decisions about whether or not to invest in and adopt these approaches.

Beyond the kinds of questions illustrated above are ***even more specific questions about the design principles that can be applied in future courseware development efforts to maximize the likelihood of obtaining positive impacts.*** In the future, those who develop courseware using funding from the foundation (or other external organizations) could be required to articulate the design principles that they use at the level of detail exemplified by Exhibit 13. Even if the developers do not test their design principles experimentally, documentation of the principles they used would at least support the identification of those principles most often associated with positive impacts. As noted, this level of detail is largely missing both in reporting to funders and in the academic research literature.[23]

In terms of the assumptions underlying the Postsecondary Success strategy and the related grant competitions, the project portfolio provided confirming evidence for some assumptions, but the jury is still out on others. The course completion data suggested that the Postsecondary Success strategy was overly optimistic about the power of circa 2010 learning technology to improve course completion rates. When mastery learning and self-pacing were used, there was a trade-off

[22] B. Means, M. Bakia, and R. Murphy, *Learning Online.* New York: Routledge, 2014.

[23] B. Means, Y. Toyama, R. Murphy, and M. Bakia, The Effectiveness of Online and Blended Learning: A Meta-analysis of the Empirical Literature, *Teachers College Record, 115*(3), 2013.

between the stringency of the mastery criterion (and hence the amount of learning of each skill or concept) and the speed of curriculum completion.

Mastery learning approaches are associated with improved developmental mathematics learning outcomes but not with improved completion rates in the absence of a major restructuring of the course sequence. Postsecondary Success courseware projects have generated large amounts of student outcome data for online mastery learning approaches in developmental and gateway mathematics courses. The CTE data from many campuses provide insight into the trade-off between the stringency of the mastery criterion (and hence the amount of learning of each skill or concept) and the speed of curriculum completion when mastery learning and self pacing are used. For many of the NCAT Changing the Equation developmental mathematics courses, completion rates actually declined after incorporation of mastery learning software even as measures on assessments of math learning rose. Although decoupling course completion from lockstep pacing enables a few students to zoom through the course material, a much larger percentage of students placed into developmental math will need more than the conventional time to reach mastery on the whole sequence of required objectives. These findings are consistent with results of earlier meta-analyses of mastery learning.[24] Mastery learning approaches involving extensive skills practice generally produce better learning outcomes, with effect estimates in the .50 to .60 range overall and with larger effects for low-ability students, but they also result in an increase in the average instructional time required on the order of 25%.[25]

To date, the Postsecondary Success courseware portfolio has provided relatively little evidence of the efficacy of more advanced or innovative learning software. The courseware that has been developed and implemented has not been groundbreaking for the most part, and the more innovative development projects have not collected evidence of learning impacts. The portfolio provides relatively little data on the effects of forms of personalizing learning other than mastery pacing—for example, changing the level of scaffolding or type of feedback based on students' prior performance on the learning system, offering alternative content addressing the same learning objective, or tailoring content to students' interests and occupational plans. More innovative types of course software, such as interactive simulations, gaming features, and virtual environments with teachable agents, represent only a small proportion of the courseware in this review.

An exception to the lack of evidence with respect to more innovative courses is the outcome data produced for hybrid implementations of MOOCs in colleges other than those where the MOOCs were developed. Outcomes from the ITHAKA S+R study of hybrid MOOCs in the University of Maryland system and the evaluation of MITx and edX hybrid MOOCs in Boston community colleges suggest that ***hybrid MOOC courses result in levels of learning similar***

---

24 John A. C. Hattie, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.* New York, NY: Routledge, 2009.

25 John A. C. Hattie, *Visible Learning.* Routledge, 2009, p. 171.

to conventional courses the first time instructors try them. Subsequent implementations might yield superior outcomes, but this assumption has yet to be tested. At the same time, the various MOOC projects surfaced a variety of implementation challenges, largely because (1) the MOOC resources were not developed with reuse by other faculty in mind and (2) early MOOC platforms lacked capabilities for giving instructors access to the kinds of formative assessment data they wanted.

The challenge of reliably achieving positive outcomes at scale remains a major issue. Some of the early Postsecondary Success courseware investments were made with the assumptions that (1) a large supply of effective technology-based courseware was being used in individual courses that was not spreading to other instructors and campuses because of market barriers and (2) if those market barriers were addressed, courseware that had been implemented successfully in one setting could be scaled broadly without loss of effectiveness.[26] The latter assumption did not bear up well in the NGLC Wave I evaluation.[27] On average, the NGLC innovations improved student course outcomes when implemented on the campus receiving the grant and had no effect on course outcomes when implemented on expansion campuses. In contrast, the Carnegie Foundation for the Advancement of Teaching has data showing positive effects for all but one of the campuses implementing its Pathways courses. The Pathways Project put tremendous effort into the design and support of its courseware and associated student supports, setting the bar for project participation at a level that limited the number of campuses it could support. The Pathways Project provides insights into what it takes to achieve consistently positive outcomes, but further work is needed to find ways to do so at lower cost to enable faster scaling. The field needs to learn much more about how to achieve reliably positive outcomes from using course models incorporating online learning at scale.

Knowledge gaps exist with respect to the ongoing costs associated with implementing courseware-based interventions. Measuring the costs of interventions involving digital courseware is an undertaking easily as complex and susceptible to bias as measuring learning outcomes. Relatively few courseware-based interventions are subjected to a systematic analysis of the costs of development, initial implementation, and ongoing implementation.[28] Although more research of this type is needed, Postsecondary Success projects have started to address this gap. NCAT has been a pioneer in working with its partners to gather data on the most significant cost drivers (instructor labor and class size). ITHAKA S+R has gathered cost element data for the course implementations it evaluated in the UMD Blended project and found that the use of MOOCs in a blended course cut instructor time in the classroom by around 50%. In general, the foundation is giving cost data more emphasis in its current courseware efforts, notably the ongoing evaluation of ALMAP.

---

[26] EDUCAUSE, *Next Generation Learning Challenges Wave 1: Building Blocks for College Completion*. Request for Proposals, October 2010.

[27] Barbara Means, Linda Shear, Ying Zheng, and Rebecca Deustcher, *Next Generation Learning Challenges Wave 1: Evaluation Final Report.* Menlo Park, CA: SRI International, 2013.

[28] B. Means, M. Bakia, and R. Murphy, *Learning Online.* Rutledge, 2014, 2014, pp. 165-177.

Courseware effectiveness research that has been done thus far tells us little about whether digital courseware contributes in the long run to degree completion, the foundation's ultimate goal. A handful of projects have collected data on student success in the course following the one in which the student experienced the intervention, but in the absence of mandated reporting, substantial bias is likely in what is reported. Moreover, none of the courseware projects reviewed here compared degree or certificate completion rates for treatment and control course sections. The closest approximation we found was the measurement of the number of college credits earned within two years of completing developmental math conducted by the Carnegie Foundation's Pathways Project. Although this absence may seem surprising in light of the foundation's emphasis on degree completion, it is understandable given the high cost of longitudinal research and the nascent state of most of the courseware interventions in the portfolio. It makes sense to invest in research on longitudinal impacts only where significant near-term impacts have been documented. Within the Postsecondary Success portfolio, those projects focused on increasing successful completion rates for developmental mathematics are closest to being ready for the evaluation of long-term impacts.

Lessons Learned

With the long-term goal of dramatically improving students' degree completion rates, the Gates Foundation's Postsecondary Success strategy seeks to understand what is required for technology applications to produce positive student impacts at scale. In this section, we highlight implications of our review for the design of future digital courseware initiatives.

In funding portfolios of projects, even modest grant amounts from high-prestige funders will attract numerous applications. The Postsecondary Success experience suggests that many organizations—not only institutions of higher education, but also nonprofit organizations and startup companies—are interested in receiving a Gates Foundation grant for developing and implementing courseware. The prestige factor and opportunities to interact with the experts and other grantee organizations the foundation convenes appear to be larger inducements than the funding amount per se in triggering grant applications. On the other hand, to the extent that the foundation is asking grantees to undertake activities that their organizations are not staffed to perform (such as systematic collection and analysis of data), attaching funding to these functions increases the likelihood that grantees can meet performance expectations. Clarity in communicating expectations for data collection and sharing is important at the stage where grantee applications are solicited.

A core insight is the need to adjust our understanding of what a courseware intervention really is. Previous efforts, not only those funded by the foundation but many others as well, make it abundantly clear that hardware and courseware alone are insufficient to produce better learning outcomes and educational attainment reliably at scale. Those technology-supported interventions that have demonstrated positive impacts on students (such as Pathways) have involved redesigning whole courses, changing instructor practices, and adapting organizational policies and allocation of time and space to align with more personalized instruction.

We also have learned that the process of spreading an innovation from one campus to others is much more challenging and complex than envisioned in earlier theories of action, such as that articulated for the first wave of NGLC grants. Interventions get adapted to fit local contexts and priorities; the key is to understand the types of adaptations that make an intervention more usable (and palatable to adopters) without sacrificing efficacy. One of the Boston community colleges, for example, had concerns about the rapid pacing of the content in the edX Python programming course. To accommodate the students, the course structure was adapted so that one week of MITx content would be delivered in two weeks at the community college using edX. Other grantees (Cuyahoga Community College, Georgia Tech) held similar views. One project leader, for example, stressed the importance of keeping the audience in mind when designing online course content and how courses developed at Harvard or MIT were generally ill suited for most community college students.

Effective innovations involving courseware are developed through multiple cycles of design, implementation, and improvement. Interventions powerful enough to move the needle significantly on student learning and course completion are not created overnight or over a single academic term. Pilot-testing in one or a few classrooms typically reveals issues that need to be addressed before an intervention can reach its full potential. The more successful projects had longer grant terms (or successive grants) and planned for multiple iterations. NCAT, for example, customarily has its partner institutions pilot-test a redesigned course before fielding it on a larger scale and evaluating impacts. Portions of the Statway and Quantway courses developed under the Pathways Project were piloted in a few classrooms, and then the courses were redesigned and implemented in a formal pilot in a larger number of sites before the third iteration that involved a large number of institutions and the formal collection and analysis of student outcome data. OLI courses, especially OLI Statistics, have gone through many cycles of iteration and modification.

The collection of appropriate data during pilot studies and initial implementations is key to improving a courseware intervention with each iteration. Iteration without data is a random walk. In addition to providing information needed to improve the course, pilot implementations help institutions learn what data they are required to collect and how to analyze student outcome data for purposes of refining their intervention. We have found that these practices are surprisingly rare at institutions of higher education, especially community colleges. Most colleges have some staff responsible for student data systems and institutional research and other staff responsible for making sure the campus IT infrastructure supports students and faculty. It is rare, however, for campuses to bring these staff together with instructors developing innovative courses to implement multiple cycles of data collection to inform course improvements. But such efforts are not unrealistic. Carnegie Foundation's Tony Bryk reported that when institutional researchers were invited into the course redesign process as codesigners, they were eager to engage in this kind of activity.

Communities of practice involving institutions adopting the same innovation can be an effective strategy for building knowledge about the requirements for effective implementation. As noted above, much remains to be learned about how to achieve consistently positive outcomes at scale, but the examples of the Pathways and NCAT/CTE projects suggest that having an active intermediary organization at the hub of a network of implementing colleges is a promising strategy. Rather than assuming that an intervention that has produced positive results on one campus will be similarly effective when implemented elsewhere, we need to document changes in context and practices at each adopting site and measure actual value added in the new context. The collective analysis of data collected from multiple implementations on different campuses is needed to build an understanding of the necessary and sufficient conditions for effective use of the courseware.[29] It takes a partnership among

[29] A. S. Bryk, L. M. Gomez, and A. Grunow, Getting Ideas into Action: Building Networked Improvement Communities in Education. *In Frontiers in Sociology of Education.,* M. Hallinan, Ed. New York, NY: Springer, pp. 127-162, 2011.

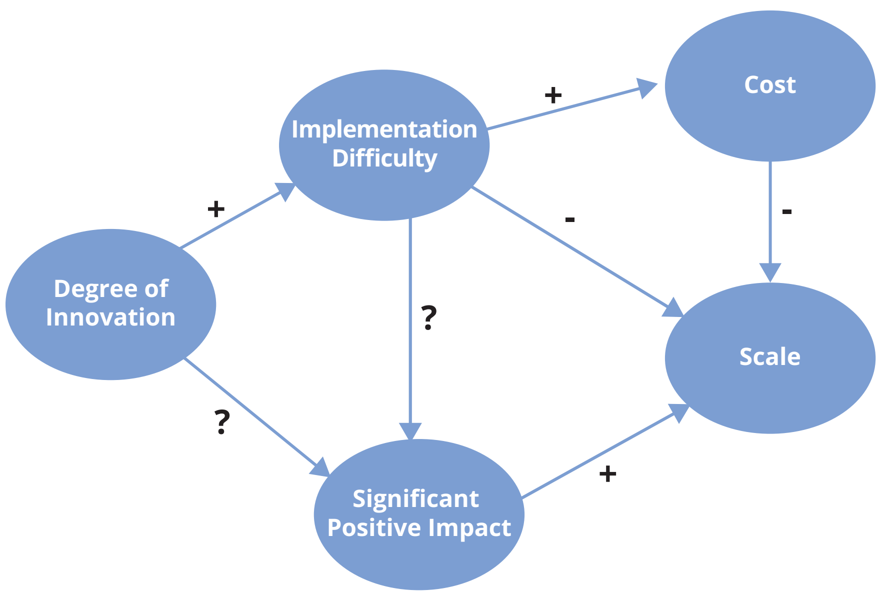
multiple organizations to collect comparable data across different course iterations and campuses and an intermediary organization to coordinate data collection and analysis.

Aggressive timelines focus grantees' efforts but do not allow for iteration and improvement. Some organizations receiving courseware grants have felt energized by tight timelines requiring getting a viable product into classroom use as quickly as possible. In some cases, however, this has meant that courseware has been fielded before thorough usability testing or the development of key features. Funders should avoid giving grants that encourage premature implementation in large numbers of classrooms. Technical difficulties can discourage instructors from ever using technology-based instructional resources in the future. Ineffective, buggy courseware can consume student time without actually returning benefits and may even reduce chances of course success. New products should be tried out on a small scale or in low-stakes situations (such as an optional course activity), with appropriate data collection and improvement cycles, prior to large-scale use.

Conceptual Framework for Courseware Investments

In this section, we offer a framework for thinking about the challenge of stimulating the development and scaling of effective technology-based courseware and the role that evaluation activities can play in addressing this challenge. We start with a depiction of the relationships among some key variables: the extent to which a technology or other educational intervention is truly innovative, the degree of change the innovation requires on the part of instructors and educational institutions, the cost of purchasing and implementing the intervention, and the scalability or demand for the innovation. These relationships are shown in Exhibit 14.

Exhibit 14. Relationship Among Courseware Implementation Strategy Factors



Education reformers regard more innovative courseware as desirable on the assumption that it has greater potential for having major, transformative impacts. But, as Exhibit 14 suggests, for any particular courseware innovation, we do not know in advance of evidence collection whether or not there will be a positive impact for students. It is safe to say, however, that more innovative technology-based interventions will generally be more difficult to implement because they require more change on the part of instructors and institutions.[30] In turn, implementation difficulty drives up costs and reduces demand.

[30] P. Blumenfeld, B. Fishman, J. S. Krajcik, R. W. Marx, and E. Soloway, Creating Usable innovations in systemic reform: Scaling-up Technology-Embedded Project-Based Science in Urban Schools. *Educational Psychologist, 35*(3), 149–164, 2000.

The provision of professional development and technical support for implementing an innovation can reduce the perceived implementation difficulty and increase the likelihood of obtaining positive impacts, but at the same time, such supports are costly, which tends to reduce demand among cost-sensitive institutions. Presumably, evidence of a significant positive impact for the innovation would make institutions and instructors more inclined to adopt it, despite the effort required. In the absence of such evidence, however, there is little incentive for adopting anything that is difficult to use or that incurs expenses.

As noted, meeting the twin goals of positive impacts and scalability is difficult in this complex landscape of conflicting priorities. The model in Exhibit 14 suggests that to obtain scalability of an innovative, potentially transformative, technology-supported instructional intervention, ***we should look for ways to reduce implementation difficulty, demonstrate positive impacts, and drive down costs.*** We believe that judicious use of a combination of new and traditional evaluation and research approaches can address these three critical factors.

An educational intervention is always about more than the software, and the extent of the provider's awareness of colleges' and instructors' other requirements for effective implementation can be a major success factor. Formative evaluation can sensitize courseware developers to aspects of their products that do not work as intended and can inform the development of explicit implementation models that can then become the focus of professional development.

Measuring the intended student outcomes as a matter of course provides the kind of data that can be used to inform revisions to both technology-based and other components of the innovation. Once an innovation has been honed to the point that it has a clear implementation model and reliably produces positive outcomes relative to competing alternatives when implemented as intended, further experimentation can identify the absolutely necessary components of the product and its implementation supports with an eye to cutting any unnecessary costs.

This process of design, implementation, data collection, and analysis takes time of course. And we must be sensitive to the fact that public release of early student outcome data (which are often disappointing) can kill enthusiasm for a promising new courseware product. We believe that it is important for organizations developing innovative learning software products to conduct their own R&D, but experience suggests that the involvement of external research partners adds value. Designing and promoting a new educational innovation are complex, all-consuming endeavors. Conceptualizing and running research and evaluation studies at the same time are challenges beyond the capacity of most organizations and certainly of the start-up operations becoming increasingly active in the learning technology space. Our review of Postsecondary Success project reports generally found stronger insights concerning implementation and better evidence with respect to impacts from organizations that were not trying to develop and promote courseware at the same time they were evaluating it.

In the past, the Bill & Melinda Gates Foundation has tried several alternative approaches to stimulating the development and scaling of courseware that could contribute to course success and college completion. Some major multiyear investments of millions of dollars have gone to organizations with strong institutional capacity and a track record of success. Other investments have gone for portfolios of many similar short-term grants at much smaller amounts (as little as $250,000) with the idea that early evidence of impact would identify the "winners" from the large field of contenders. The first strategy is risky to the extent that foundation staff cannot be sure their few very large investments are going for courseware that will be effective and will scale extensively. The second strategy entails the risks of promoting immature products with little prospect of near-term success for instructors and institutions and of judging courseware concepts prematurely on the basis of their first implementation with a single cohort of students.

The foundation's latest Postsecondary Success courseware grant competition, the Next Generation Courseware Challenge, reflects a new phased-funding approach that seeks to strike a balance between these extremes and that recognizes the need to iterate for improvement. The foundation is funding seven cross-functional teams to refine, implement, and collect outcome data for exemplary courseware for high-enrollment lower division courses. Each project will receive a 36-month grant of $1 million to $5 million for this work, which is to include collecting the kinds of data needed to identify those features and components of the courseware and its associated implementation model that are necessary and sufficient for obtaining positive outcomes.

Evaluation Approaches Suitable for Different Kinds of Courseware Investments

Our framework for thinking about data collection and evaluation activities appropriate for different stages and types of courseware projects is shown in Exhibit 15. The framework suggests that the objective of a courseware investment (the driving question it seeks to answer) determines the nature of the needed data collection activities, the audience for evaluation findings, and the relative size of the investment in evaluation and data collection relative to that in courseware and technology tool design and development. (More circles represent greater investment.) Brief definitions for the various research and evaluation approaches are found in the sidebar; more extended descriptions and examples of their application to technology-supported learning interventions can be found in the report *Expanding Evidence Approaches for Learning in a Digital World.*[31] That report also discusses how the traditional association between a product's level of maturity and the size of its user base has been disrupted by Internet-based educational resources, courses, and applications. Many technology developers launch early-stage "minimally viable products" on a broad scale and mine data from their large user base to obtain insights into how to improve their product.

This "high-tech" approach works well for commercial consumer products, but the situation is more complicated for courseware intended for publicly funded education institutions. Regardless of the purchase price, courseware that entails

[31] U.S. Department of Education, 2013.

Exhibit 15. Evaluation Framework for Courseware Research and Development

Goal	Evaluation Methods	Intended Outcome	Primary Audience	Development Investment	Evaluation Investment
Design and develop potentially powerful courseware	Draw on learning science research → Usability testing → Learning analytics → Design research → A/B testing →	Smarter design Improvements to user interface Usage patterns from a large number of learners Insight into fit with constraints of settings and institutions as well as instructor concerns and preferences Incremental improvements to product as measurable using product-internal data **Viable Product**	Internal	●●●	●
Understand the "wrap around" needed to get positive results using the courseware	Implementation studies → Efficacy studies →	Understanding of conditions and supports needed to implement the courseware/tool in a way that produces student outcomes Evidence that the courseware/tool can produce positive outcomes when implemented with strong support **Proof of Concept & Recommended Implementation Model**	Internal	●●	●●
Determine consistency of positive outcomes when implemented at scale	Effectiveness study (RCT) or multiple quasi-experiments with statistical controls) →	**Evidence of Reliably Positive Impacts in Realistic Settings**	External	●●	●●●
Determine whether the courseware enhances long-term learning and educational outcomes	Longitudinal effectiveness study →	**Evidence of Impact on Longer-term Outcomes**	External		●●●
Access cost efficiency	Productivity studies combining cost and effectiveness →	**Cost Effectiveness Data**	External		●●

Definitions or Research and Development Approaches

Minimally viable products (MVPs) are a strategy for accelerating the development of a product to shorten the time to market. This strategy consists of an iterative process of idea generation, prototyping, data collection, analysis, and learning. An early-stage MVP is launched online with the goal of obtaining as many users as possible, whose interactions with the product will provide information and feedback that can be used to refine the product, with many iterations over time.

Usability testing is applied with new products to ascertain the ease with which representative users can execute typical tasks using the product. Users are observed to ascertain whether they can complete the tasks successfully, the amount of time they require to complete tasks, and their level of satisfaction or dissatisfaction with the product. For convenience and to support recording of the details of user interactions with the product, usability testing is often conducted within a formal laboratory setting rather than in the contexts in which the product will eventually be used.

A/B testing allows for systematic comparison of particular features of an online system to support product design decisions. Two randomly assigned groups of users receive contrasting versions of the product that vary in a defined way: One version has feature A and the other has feature B, but they are otherwise identical. Because it is a type of RCT, differences in user behaviors with the product can be attributed to the feature being manipulated.

Educational data mining involves the application of techniques from statistics, machine learning and computer science to analyze data collected during teaching and learning. It can be employed in combination with A/B testing.

Design research involves both the design and the study of educational through a partnership between designers, researchers, and representatives of the intended population of end users (e.g., instructors and/or students). Both qualitative and quantitative research approaches may be applied, typically with a relatively small number of users. Iterative cycles of testing and refinement are used to improve the product and to generate "design principles" with broader applicability.

Implementation studies examine the way in which a particular intervention involving a digital learning resource is used in different settings, with the goal of identifying barriers and facilitators that influence whether and how teachers and students use the resource. Such studies may suggest ways in which the technology itself can be improved as well as local conditions and practices that increase the probability of an effective implementation.

Quasi-experimental designs employ a treatment and a comparison group without random assignment to groups. Instead, learner characteristics or a pretest measure related to the study's outcome are measured prior to treatment and statistical controls are applied to correct for any pre-existing group differences.

Randomized controlled trials (RCTs) employ random assignment of study participants to a treatment group that receives the intervention being studied or a control group that experiences "business as usual." Random assignment of adequately large samples of study participants insures that the only difference between participants in the two groups is whether or not they received the treatment. Thus, an RCT design is intended to guarantee that any difference in outcomes after treatment can be attributed to the treatment, thus ruling out competing explanations for those differences.

Efficacy studies employ an RCT design within a particular setting or small number of sites in which the supports needed to make the treatment effective are present. The goal of an efficacy study is to test whether an intervention can produce a desired effect under ideal conditions ("proof of concept").

Effectiveness studies employ an RCT or controlled quasi-experimental design in a large number of sites with diverse characteristics. The goal is to determine whether or not a desired effect is found across a range of real-world conditions.

Productivity studies in education combine information about the costs of an educational intervention compared to a specified alternative approach with information about the outcomes produced by the two alternatives. Educational productivity studies provide decision makers with information about how much improvement in terms of outcomes they could obtain for a unit of increased investment (in terms of dollars or time).

Longitudinal effectiveness studies measure outcomes for the same groups of students at two or more points in time.

significant costs (in instructor training or infrastructure) for the implementing institution or that students will be compelled to use for a required course or program should have evidence of effectiveness before being scaled broadly.

This caveat about scaling unproven education products is not to say that courseware developers should not think about the scalability of their products from the earliest stage of design and should not lay the groundwork for wide-scale adoption. The Carnegie Foundation for the Advancement of Teaching, for example, recruited cooperating colleges for its Pathways Project and obtained approval of its Statway course content for college credit from the American Statistical Association before it ever began having its partner institutions implement the course. In the same vein, recent articulations of design-based implementation research advocate for designing for scale from the outset, even though initial implementations are typically small in scale.[32]

[32] N. Sabelli and C. Dede. Empowering design-based implementation research: The need for infrastructure, *National Society for the Study of Education, 112(*2), 2013, pp. 464-480.

Recommendations for Courseware Investments

Drawing on the lessons learned and knowledge gaps identified here as well from as our years of experience studying learning technology, we offer recommendations for consideration by the foundation and other organizations supporting learning technology R&D.

A strong rationale remains for investing in high-quality courseware for lower division courses designed with reuse in mind as well as in research on effective strategies for scaling the most effective of them. Despite the growing availability of open educational resources and open-source learning platforms, most faculty members continue to develop their own idiosyncratic courses and learning assessments. Developmental and lower division gateway courses are roughly similar across many institutions. Nevertheless, the tradition of individual faculty control of course content and pedagogy remains strong, particularly at four-year colleges and universities. Postsecondary Success has been sensitive to the fact that most college instructors expect to design their own courses and are unaware of many of the best course designs and digital learning assets in their field. A number of Postsecondary Success initiatives have attempted to address this market barrier by providing grants to enable organizations with courseware regarded as successful to spread it to other campuses. These efforts have been successful in finding campuses willing to try courseware developed at another institution; but issues of fit have emerged, and instructors' desire to modify courses to fit their own programs, preferences, and students have been stymied because most of the courseware is not easily modifiable. The MOOC platforms available in 2013 did not support modular courseware, for example. Faculty members could incorporate the archived version of a whole MOOC into their course but could not select and reorder modules of the MOOC to fit their own teaching approach or to mesh with other course material. Much could be learned from efforts to design modularized courseware that allows for easy addition of additional resources and for adding, dropping, and resequencing learning and assessment modules.

The maturity of instructional courseware and its prior evidence of effectiveness should be considered when making the trade-off between breadth and depth of investments. Under its Postsecondary Success strategy, the foundation has made both sizable investments in individual courses and instructional systems (e.g., Pathways, NROC/DevMath) and relatively small grants to organizations applying similar approaches and technologies toward a defined educational challenge (NGLC Wave I, the Dev MOOC portfolio, and ALMAP). We conclude that there is value in both types of investment, provided certain conditions are met. Targeted, challenge-based grant programs can accelerate knowledge building for the field when a clear, common objective is specified for the program and independent formative evaluation activities are funded concurrently. By comparing the implementation issues, design features, and early outcomes for the range of approaches taken by different grantees addressing the same educational challenge, a funder can gain insight into more and less effective approaches. Achieving major impact from a particular courseware intervention at scale, on the other hand, is likely to require greater investment in a particular organization to enable the development of capacity for iterative design and testing and for scaling to large numbers of institutions. Such investments should be accompanied by the collection of evidence of impact more rigorous than required for the smaller scale investments within a portfolio of like projects. We recommend limiting major single-institution investments to innovations with credible evidence of positive impact and with a strong team and organizational capacity to take on the serious challenges posed by efforts to scale while maintaining or enhancing effectiveness.

Courseware funders should put more emphasis on third-party, independent evaluations of impact. The meta-analysis reported here was dependent for the most part on grantees' own reports of comparative outcome data for which there was no information about the comparability of students in the treatment and comparison classes. This weakness is attributable to the fact that only a minority of the Postsecondary Success investments have been accompanied by an investment in appropriately rigorous evaluation by an objective third party. Courseware design, development, and implementation are complex undertakings as is the collection and analysis of student outcome data. Many project teams have not had the organizational capacity to perform all these functions well at the same time. In most cases, data collection and analysis have suffered when project personnel were pressured to complete course development and be ready to implement at the start of the designated academic term. Moreover, when projects did combine evaluation activities with course design and development, there was an understandable tendency to pick the outcome measures that made the course innovation look best. Third-party evaluators can offer increased objectivity, use consistent measures and methods, and take responsibility for synthesizing findings across individual projects.

Philanthropy can play an important role in promoting iterative design cycles and standards for measuring the effectiveness of innovative instructional approaches incorporating digital learning. Given the absence of norms promoting systematic evaluation of instructional innovations in many higher education institutions, significant support for such practices will be necessary. Funders can specify procedures for product iteration and evaluation to enhance the likelihood that the courseware they invest in will produce positive benefits for students and higher education institutions. In addition, organizations funding the development of innovative courses and courseware may want to consider having an outside organization serve as an intermediary and technical assistance provider to bolster the evaluation expertise available to higher education institutions, increase objectivity, and obtain consistent outcome measures across projects.

Funders should take a phased approach to supporting courseware innovations, with later stages of funding dependent on demonstrated capacity to collect data that can inform improvement. Designing grant programs with stages of funding permits the encouragement of new ideas reflected in innovative designs while reducing the risk of implementing an intervention that may prove ineffective when tried on a wide scale. It does not make sense to fund the widespread scaling of an innovation that has no evidence of effectiveness.

Funding decisions and evaluation activities should be tightly coupled. Ideally, the evaluation criteria for grants that will be evaluated should be articulated at the time funding competitions are announced. Grantee organizations should understand what is required of them to produce data for the funding organization and the evaluation To extend the knowledge base in a highly evolving field such as online learning, it is essential that researchers work toward cohesion and consistency for collecting comparable data that can be used to make evidence-based claims about improved student learning outcomes and increased access to education.

Funders should consider market pull mechanisms, such as prize competitions or payment for success, as an alternative strategy for increasing the supply and visibility of effective courseware. Most government and private funding to promote educational innovation involves “push programs” that pay for R&D inputs. But interest is increasing in the alternative of “pull programs” that provide funding and other incentives for successful R&D outcomes.[33] An example of using prizes to incentivize technology development was the competition for automated essay scoring engines run by Kaggle, which garnered over 200 entries, including a number that out performed the top-selling commercial essay scoring software. An example of payment for success was the original funding mechanism for the Florida Virtual School, which provided payment for successful course completions rather than for course enrollments. Pull strategies can be very cost effective by stimulating external investments of time and money, but they require clear delineation of the intended outcome, criteria for success, and the process by which products will be judged. Like the other recommendations above, they entail the integration of evaluation into the R&D investment process.

Understanding of how to build and implement effective courseware could be facilitated by grant-making targeted on design principles. Most philanthropy around courseware and other learning technology is not designed to produce generalizable knowledge for the field. Grants go to organizations to develop courseware or broader interventions incorporating courseware with the goal of finding something that works. When the effectiveness of the intervention is measured empirically and objectively, it is the impact of the intervention as a whole, rather than the effects of particular course design features, that is being measured. Given the short shelf life of most individual courses and of start-up organizations with early-stage learning technology applications, funding organizations might want to consider R&D grant programs that identify and test design and implementation principles explicitly.

[33] We are indebted to Thomas Kalil at the Office of Science and Technology Policy for bringing to our attention the applicability of this strategy to learning software.

A portion of the investment in evaluating courseware and related technology tools should be devoted to examining longer term impacts with implications for degree completion. Longitudinal studies take time to execute and can be resource intensive. It makes little sense to try to study the long-term impacts of every courseware intervention. But those interventions that have received extensive funding and that have demonstrated large and dramatically positive near-term student outcomes warrant this kind of study. This need is especially appropriate for the Bill & Melinda Gates Foundation's Postsecondary Success strategy, given its stated mission of dramatically improving college completion rates for underserved populations.

References

Blumenfeld P., Fishman, B., Krajcik, J. S., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling-up technology-embedded project-based science in urban schools. *Educational Psychologist, 35*(3), 149–164.

Bork, R. H., & Rucks-Ahidiana, Z. (2012). *Virtual courses and tangible expectations: An analysis of student and instructor opinions of online courses.* Paper presented at the Annual meeting of the American Educational Research Association.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2012, May 22). *Interactive learning online at public universities: Evidence from randomized trials.* Available at www.sr.ithaka.org.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127-162). New York, NY: Springer.

Carnegie Foundation for the Advancement of Teaching. (2013, October 15). *Quantway and Statway: Creating new pathways for advancing the success of community college students*. Annual report to the Bill & Melinda Gates Foundation. Stanford, CA: Author.

Clyburn, G. (2013). Improving on the American Dream: Mathematics pathways to student success, *Change: The Magazine of Higher Learning, 45*(5), 15–23.

Figlio, D. N., Rush, M., & Yin, L. (2010). Is it live or is it Internet? *Experimental estimates of the effects of online instruction on student learning*. NBER working paper 16089. Cambridge, MA: National Bureau of Economic Research.

Frank, K. A., Duong, M. Q., Maroulis, S., & Kelcey. B. (2011). *Quantifying discourse about causal inferences from randomized experiments and observational studies in social science research.* Austin, TX: University of Texas.

Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education,* 10(3), 157–172. doi:10.1016/j.iheduc.2007.04.001

Jaggers, S. S. (2011). *Online learning: Does it help low-income and underprepared students?* CCRC Brief, No. 52, March 2011. New York: Columbia Teachers College, Community College Research Center.

Lovett, M., Meyer, O., & Thille, C. (May, 2008). The Open Learning Initiative: Measuring the effectiveness of the OLI statistics course in accelerating student learning. *Journal of Interactive Media in Education.* Available from http://jime.open.ac.uk/jime/article/viewArticle/2008-14/351

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

Means, B., Shear, L., Zheng, Y., & Deutscher, R. (2013). *Next Generation Learning Challenges: Evaluation of Wave I*. Menlo Park, CA: SRI International.

Sabelli, N., & Dede, C. (2013). Empowering design-based implementation research: The need for infrastructure. *National Society for the Study of Education, 112*(2), pp. 464–480.

Twigg, Carol A. (2012, January) *Improving learning and reducing costs: Project outcomes from Changing the Equation.* Available at http://www.thencat.org/Mathematics/CTE/CTE_Lessons.html.

U.S. Department of Education, Office of Educational Technology. (2013). *Expanding evidence approaches for learning in a digital world,* Washington, DC. Available at http://www.ed.gov/edblogs/technology/files/2013/02/Expanding-Evidence-Approaches.pdf.

Wise, A. F., & O'Neill, K. (2009). Beyond more versus less. In S. Tobias and T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* NY: Routledge.

Xu, D., & Jaggers, S. S. (2011a). *Online and hybrid course enrollment and performance in Washington State community and technical colleges.* CCRC Working Paper No. 31, March 2011. New York: Columbia Teachers College, Community College Research Center.

Xu, D., & Jaggers, S. S. (2011b). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis, 33,* 360.

Appendix A: Project Leader Interview Highlights

The pages that follow contain highlights from interviews conducted between March and April of 2014 with the following principal investigators from Postsecondary Success Courseware projects:

Interview with PAUL M.A. BAKER, Georgia Institute of Technology

English Composition, Psychology, and Physics MOOCs

SRI: What do you consider to be the greatest contributions of your project?

One of the most interesting outcomes is the increased interest among faculty in exploring new models of teaching, specifically, in thinking about what works and what doesn't work. Originally when we started C21U, in faculty discussions on teaching innovations, about a third were interested in new platforms/approaches, a third were neutral, and another third were somewhat hostile. It's ironic that for a liberal institution, the academy, that people are pretty conservative in their approach. But the rollout of the Gates MOOC has brought new interest in teaching to faculty. In fact we would now (2 years later) say that 50% of faculty are enthusiastic or at least very interested in new learning approaches, and the balance neutral but open to new ideas.

Partly as a result of our experience with developing MOOCs, as well as our ongoing monitoring of innovation in in higher education, we now do some advising for the University System of Georgia, as well as with some other systems. We were lucky that when the grant solicitation came out, we had been thinking about the topic pretty deeply. Even so, we had about three weeks to develop proposals, we submitted 5 MOOC proposals and were fortunate too get three funded. This was all very rapid, and we were developing our thinking on the fly. The fact that they worked at all is amazing, much less turned out to be very successful.

SRI: What informed your design decisions for your courseware? What would you do differently to improve the overall design?

The way the grant was written it really was optimal to go with an established provider. The whole program had to start quickly. The awards were announced in October for courses staring the following spring. The physics course started late, mostly due to complexity in getting release for use of course material in an open context, as strongly suggested by the RFP. The Composition MOOC ran into problems due to trying to make peer grading/evaluation work beyond its capacity – the humanities are much harder to evaluate than objective science subjects. The MOOCS were developmental and designed for Gen Education, and Coursera provided the platform, Georgia Tech faculty developed the content. There was an iterative process during the course, and things were being designed as it went along. In theory courses are designed to be used multiple times, but not forever, and not to run without faculty involvement. The degree to which there's alteration and improvement matters. The same course can't just keep running without tweaking, especially in fields in which developments regularly occur. We were seriously limited by the Coursera platform constraints, and by the fact that content was developed and formatted as the course was going along. They facilitated production but were not especially creative about platform capacity or allowing hooks to outside software with more flexible capacities.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

Developing, ramping up and maintaining a MOOC has significant costs if done properly. If [you're] using a big platform, as opposed to a "basement computer approach" it costs between 100 and 250K the first time, including the instructor, management and ancillary support (e.g. legal advice for content, instructional design, admin, etc.). Developing a new MOOC can take the resource equivalent of two or maybe even three courses. First there's developing it, then second is running it, even with teaching assistants. The first time is expensive, and even the second time can run about 50% of the cost. Even the third time you run it, the costs can go up again because you have to update content. It can not be stressed enough how important good teaching assistants are. In our case they kept the forums running and provided support to students taking the classes, especially during times when the professors were not online. Overall, the Gates grants covered about half the cost of producing and running the MOOC, and the balance was covered by center and university overhead.

SRI: What student learning outcomes resulted from your project?

To some extent that depends on what you consider to be meaningful outcomes. What can you really answer? Does a MOOC work? Yes, if the metric is how many people signed up for a course – this is a significant measure of interest in a topic. But how many people are actually engaged (participating) with the course versus indicating an interest? We found that the higher the commitment that students have, the higher the likelihood of completion. As an instructional tool, MOOCs have significant utility. What makes it work is engagement, having a community and being involved. The community forums made it work – they were mostly asynchronous. The big questions about MOOC outcomes are still being explored: What did you learn about pedagogy, new forms of instructions, or grading and interaction? MOOCs are a useful instructional tool, and represent a significant addition to the instructional toolkit. The fact that 30,000 signed up and 300 finished is irrelevant. What matters is the output. Was our MOOC working with a specific population with special needs? We don't know. And in terms of efficacy, we weren't set up in a way that we could measure it directly. In broad policy terms the answer is more "do the participants feel like they received benefit form participating?" Mere exposure to a subject might improve subsequent attempts. For individuals in developing countries, MOOCs might represent a critical exposure to higher education that they might not otherwise have had. So the outcomes are to some extent dependent on what questions you ask.

SRI: What are your plans for sharing or furthering the work from this project?

First there's the question of "Are there Georgia Tech students and are you doing research?" If you're working with [Georgia Tech] students, you're bound by FERPA [Family Educational Rights and Privacy Act]. If our students are enrolled in the course, we have to create a special case of it that protects the data from public release. The biggest issue in this project had a lot to do with IRB [Institutional Review Board] because we have the first semester-length MOOCs. The IRB here is very engaged, especially given the newness of the whole MOOC approach. I have an email thread that was 40 emails long about a single survey item. It was one of the most onerous parts of the project. That said, we continue to leverage the interest and attention drawn by the MOOCs. Georgia Tech intends to have more than 40 MOOCs completed, running or in production by next year. We have an online Masters in Computer Science that is entirely MOOC based. We have presented finding of the Gates MOOCs and other MOOCs we are running at a number of conferences and have several papers in publication both from C21U as well as from other MOOC instructors. Finally, we have ongoing projects exploring new models of learning in conjunction with the University System of Georgia, as well as other partners. We have submitted a grant application to the Department of Labor, in part as a result of our work exploring competency- based and alternative learning approaches, and we are in the process of exploring the development of a network of higher education innovators.

Interview with DENISE COMER, Duke University

English Composition MOOC

SRI: What do you consider to be the greatest contributions of your project?

The fact that we successfully developed a MOOC that uses the most effective practices in writing pedagogy – the fact that it worked. This course was the first Introductory Composition MOOC of its kind. There was a real sense of online community among learners. Students had the opportunity to crowd-source an annotated bibliography and volunteer to participate in a Google Hangout writing workshop either in small groups, large groups, or one-on-one. Students also had a global audience, which is important because people don't necessarily know how to communicate across cultures. This aspect was both challenging and rewarding. Some students contributed an Op-Ed piece for the course, and had their work published around the world.

SRI: What informed your design decisions for your courseware? What would you do differently to improve the overall design?

We worked with Coursera to develop the course. It was a challenge to transform dialogic pedagogy into Power Points and videos. We ended up recording and editing 77 videos for the course. Students interacted by having discussions in the online forums, mostly by asking questions and receiving answers. They also used Meetups and Google Hangouts to collaborate. In future courses, I'd like to know how to approach a course that has a lot of linguistic diversity among students.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

We worked very hard to use pedagogical best practices that are standard in the writing discipline and for online instruction. That was our implementation model. Within the course, we relied extensively on peer assessment. We used highly structured rubrics to provide feedback toward revision on drafts and also structured grading scales to evaluate final versions. After the course, we conducted a novice-to-expert rating comparison.

Another important aspect was that students could pace themselves. They could choose which pieces of the course they wanted to complete. The MOOC also made visible the importance of differentiated learning. Students learn for different purposes, in different contexts, and have different outcomes and motivation. In college, learning is top-down, and the instructor hands the materials out and dictates the learning objectives. In a MOOC, the learners define for themselves what they want to do. For example, maybe they just wanted to make a friend with whom they can exchange writing.

SRI: What student learning outcomes resulted from your project?

We did track learning outcomes. We held an intensive portfolio rating session with an IRB-approved portfolio sample. If students from our own institution participated, we didn't know about it. We knew how many students participated in discussion forums, and how many submitted writing drafts. The completion rate was 1,289 from a total enrollment of approximately 64,000. It's difficult to get a representative sample in a MOOC because you can't tease the students apart. Students completed self-reflection quizzes and indicated through self-assessment their progress toward learning outcomes. The data available from the MOOC is so big that it can seem unwieldy and inscrutable; it emphasizes the importance of cross-disciplinary collaboration in understanding MOOC data.

SRI: What are your plans for sharing or furthering the work from this project?

We're doing a qualitative analysis with NVIVO to analyze peer engagement and interaction. We're asking the question – What is the impact of peer-to-peer interaction on learning in a MOOC? What is the motivation of a post? We're partnering with an instructor who teaches chemistry for this project, which is funded by Athabasca and Duke. I'm writing some papers on this work – some journal articles and a book chapter. The MOOC will be run twice this year, in spring and fall. It will stay on Coursera's list of offerings.

Interview with REBECCA GRIFFITHS & MATTHEW CHINGOS, ITHAKA S+R

University of Maryland MOOC Blended Course Project

SRI: What do you consider to be the greatest contributions of your project?

We think we are seeing from the instructor side that there were benefits for them. We interviewed 18 instructors, and the large majority (15 of 18) said they would do the MOOC again and would recommend this model to their colleagues (13 of 18). Common themes from instructor interviews were: Opportunities they see to provide students with different perspectives, to augment their courses with material outside their expertise, flipping the classroom, and saving time in some cases. Some instructors expressed concern about the amount of time it took to prepare the course. Six (all adjuncts) said it saved time, but others spent lots of summer time viewing all the [MOOC] lectures. Many of those said it might save time in the future [if they re-used the same MOOC].

SRI: Did you specify a specific implementation model for the hybrid courses?

No. We wanted to leave it to the instructors. Instructors came up with some creative ways to use the MOOCs. One used it to train students how to consume an online course. It was a learning community. With three courses built around the Duke To Think & Reason course. They met with an undergraduate 3 times a week and went through the process of figuring out how to glean information from the online materials. The instructor thought it worked very well, and that having a common set of literary experiences would be useful for her students.

SRI: From your experience, how important is it to have a face-to-face component that supplements online learning?

Instructors felt that having a face-to-face component was really critical. Many faculty said it was hard to make sure students were doing the online work. Coursera doesn't have a dashboard that works well for small classes. Its dashboard is designed for thousands of students, not set up to show you how each student in a 30-student class is doing. Instructors did use the dashboards for the OLI [Open Learning Initiative Statistics course] and Pearson tests. Instructors of Coursera classes wished they had had better tools [for looking at student progress].

SRI: What design recommendations would you give to technology developers for creating online learning resources for use in blended college courses?

There were a lot of technical issues using Coursera. It felt like two different courses to students because it was not integrated with the campus's course management system. If MOOC platforms want to license content to institutions, they need to make it easier for faculty to integrate it with their courses and more user friendly to students. It needs to be more of a plug-in to the LMS [learning management system] like MyMathLabs is configured. Then a dashboard for instructors to monitor individual-level assessments [is needed]. Instructors wanted to use the peer evaluation function from MOOCs, but it doesn't work well with small groups of students.

SRI: What student learning outcomes resulted from your project?

We can't talk about anything except instructor perceptions until the student data are analyzed. The Genetics course at UMES [University of Maryland Eastern Shore] had an instructor who thought students would be more engaged by MOOC materials than by publisher materials. He didn't think he saw any increased engagement.

SRI: Were you able to analyze student-level log data for outcomes?

Some logs are more useful than others. Coursera logs don't capture many important student activities. Pearson logs are surprisingly difficult to interpret. They seem inaccurate since the data don't correlate with instructor descriptions of what went on. We haven't got OLI logs yet.

SRI: What are your plans for sharing or furthering the work from this project?

We will do a SXSWedu panel with some faculty members and a public report in May or June. [See http://www.sr.ithaka.org/research-publications/Interactive-Online-Learning-on-Campus.] Some of the faculty members will share their stories. One is presenting at a Coursera conference. Someone from USM [University System of Maryland] is doing a workshop at Emerging Techs in April. We'll also do a plenary session at the AASCU [American Association of State Colleges and Universities] provost meeting.

Interview with KAY HALASEK, Ohio State University

English Composition MOOC

SRI: What do you consider to be the greatest contributions of your project?

Some of our most compelling outcomes had to do with scaling a course. Can the teaching of writing be scaled? There are people offhand in the field who would say "no," that the teaching of writing may not scale if it's informed by traditional pedagogy. But the teaching of writing can scale if we conceive differently of the role of student and instructors in a course.

SRI: What informed your design decisions for your courseware? What would you do differently to improve the overall design?

The course was designed around a system of peer review. I'm seeing anecdotally the same kinds of detailed commentary – a fairly specific and systematic approach to peer review that articulates particular practices. It's a criteria-based assessment. It's both focused and specific, and a critical element.

If we didn't have [criteria-based assessment] in the MOOC, then we were less likely to get peer review that was robust. Then we can start looking at the points and attributes that the reviewers found. But looking at data this way takes resources, but we need these kinds of analysis before we can start asking questions.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

One of the major elements we created was an anonymous peer review system called WEx. We hired a computer programmer to write a system that houses all the assignments, as well as the peer reviews. Each of the writers gave a helpfulness score for the peer reviewer. We ran it in the MOOC and as a pilot across three sections with about 70 students in our second-year writing course. It was really important to have peer review functionality in the MOOC, where students submit a paper, get peer review, and then reflect. But it could be improved.

SRI: What student learning outcomes resulted from your project?

There wasn't a control group, so we couldn't do a comparative analysis. Right now, we don't have any quantitative data that peer review in a MOOC works. We gave out a voluntary disenrollment survey that Coursera developed. The great majority of students were dropping out because they were just dropping in to see what it was like in the first place, or there were forces in their personal life. It wasn't about the course structure of the MOOC.

There are all these data but little research support. We're sitting on this massive amount of information, but we'd have to hire someone who could look at the data. We're interested in seeing if there were any correlations between writing quality and reviews. You'd think that if a student were a good writer, they'd be a good reviewer, but that's not necessarily the case. There doesn't seem to be a correlation. Most of the MOOC cohort was older and already had bachelors degrees.

SRI: What are your plans for sharing or furthering the work from this project?

We're continuing to look at peer review as part of our extended research, and examining the role of the discussion forum and community online. We're interested in looking at natural language processing – we applied for additional funding for that but didn't get it. We ran this course in Spring 2013. We made some revisions to the WEx system and anticipate running a 6-8 week course in August.

Interview with DAVID HARRIS, OpenStax College/Rice University

OpenStax

SRI: Could you describe the background of your courseware project?

Connexions started in 1999 with a Signals Process textbok, we wanted it to be available free to the developing word. There really wasn't a platform for open learning resources, so Dr. Richard Baraniuk, being a maverick, created one for himself. To dates there are over 20,000 learning modules and over 1 million users visit the site each month.. Four years ago looked at the analytics. It was obvious we were reaching students, but not the at-risk students we were targeting. OER hadn't moved into mainstream. How do we close the chasm? The purpose of OpenStax was to create the highest quality open education resources in the world, using peer review and working with content developers. In just over two years our books are in use at over 800 institutions and we estimate over 130,000 students in courses will use an OpenStax College book this fall.

SRI: What is the design process for developing OpenStax textbooks?

If I look at most college courses, there is 80-90% overlap, and probably 10% differentiation. Scope and sequence of content is key. The intellectual effort involved in making a textbook is traditional – you can't automate that. It's a four-stage process. We work with Ph.D.'s that write the content, then it gets peer reviewed, then we work with a content developer, then it's reviewed again for copy editing. After a textbook is published it gets reviewed again. There's a panel that adjudicates any reported errors.

As far as getting books in students' hands, it starts with faculty. It's the faculty that drives the content – they want to see certain topics in textbooks that reflect their own knowledge and expertise. People often blame publishers, but it's the market that drives the content. We have multiple modalities to engage the market. We don't have a sales force so we deal with institutions directly. We have over 15 ecosystem partners, including WebAssign and Wiley, that provide services around our content.

SRI: How would you describe the cost savings benefit of your product?

We only consider verified student enrollment to calculate student savings. With the data, we estimated a savings of $11.8 million. What's amazing is that we had this level of impact in the market without a sales force. We estimate on the conservative side. Right now an introductory Physics textbook is $280. In reality, a certain percentage of students will buy used books, a certain percentage will buy new, and others won't buy the book at all. They discount the average saving down to $98 Our site analytics show that that our projects have been viewed millions of times and downloaded nearly a million times.

SRI: What are your plans for sharing or furthering this work?

We're working on OpenStax Tutor courseware. It has cognitive science principals and machine learning algorithms as a design base.. College faculty are skeptical by nature, that was their training so you can only introduce so much change at a time. But if you loose that skepticism slowly over time, you will change the market. That's the big lesson. What we [OpenStax] produced is not pedagogical reform, but leveraging openness for using the texts in an intelligent way.

If the publishers are the enemy in this, they are counting on the open community to be radical. And they're counting on foundations to support bold initiatives. That's their calculation. They're really good at being market experts and getting products into the market. They're counting on open resources to be overly progressive. You can't give [institutions] a full solution. You have to give them an 80-90% solution and let them work with it. Do less but impact more. That's a lot better than making lots of changes to materials but few students use them.

Interview with LAURIE HARRISON, University of Toronto

Psychology and Statistics MOOCs

SRI: What were the characteristics of your Courseware project?

There were two MOOCs that were really different. The [faculty member] in the Psychology MOOC is an award winning instructor. He was already teaching at scale, and already had tools and methods developed for online learning. For this project he gathered information about three existing tools and scaled them up even further to engage students in active learning processes. Because he's in psychology, he used a profile questionnaire based on the Big 5 personality matrix. He explored these as indicators, as data to learn about student completion in the MOOC. In the statistics MOOC, we found most of the students already had graduate degrees. I know [the Gates Foundation] has a mission of serving underrepresented students, but it turned out those were not most of the students in the course. Even though this is a statistics course—the instructor is parsing written assignments in the MOOC to evaluate students' higher order thinking as part of the research. They are analyzing students' written explanations based on established models of statistical literacy. The first pass [in analysis] was to learn who shows up for the course, etc. Now it's more focused research… also using an inverted classroom. The population that they're sampling in the flipped course is more controlled. [The faculty] did really well with borrowing what we already knew about online learning to improve the quality of their MOOCs.

SRI: What informed your design decisions for your courseware?

We were one of the earlier partners that used Coursera. Then we started an edX cohort of instructors as a group, so we held a workshop day on MOOC development. There were also additional sessions on assessment design, peer assessment, and how to write rubrics. It bootstrapped a lot of other things, and thinking of MOOCs as part of a larger capacity for developing learning technologies.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

There is lots of iteration, now the MOOC is playing back into their classes on campus. We have a community of practice that holds roundtables on whatever is going on with the MOOCs, whatever new challenges the faculty are facing. A community of MOOC instructors now that there's enough of a critical mass– they have interdisciplinary connections with teams from engineering, computer science, business school, aboriginal studies, there will be a MOOC research symposium. There's enough interest in MOOCs to have an in-house conference, and we did invite attendees from beyond our campus.

SRI: What student outcomes resulted from your project?

There were really different kinds of datasets that were associated with the project. We used seed money from [the Gates Foundation] to put together an institutional database of MOOC data. It's a real challenge to make sense of data-dumps from Coursera. We have a powerful dataset, but the average researcher couldn't make sense of it without someone who's a data wrangler. We do have some student-level data from a spin-off from [the Gates Foundation project] where we're looking at inverted classrooms and exploring different levels of student use of a course for review or remediation. Those inverted classrooms are U of T [University of Toronto] students, so we do have a mapping between their student ID and the anonymized system data. That data is still being analyzed.

SRI: What are your plans for sharing or furthering the work from this project?

Both instructors continued with their initiatives, some of it is research, some teaching practice. Not sure if [the Gates Foundation project-funded] instructors will continue with the new MOOC production. Some of the outcomes may not take the form of MOOCs. They have folks that want to keep going with the inverted model, and some others are launching open, public MOOCs. Publications will follow at some point—academics don't want to part with incomplete/wrong data.

Interview with LAURA KALBAUGH, Wake Technical Community College

Pre-Algebra MOOC

SRI: What do you consider to be the greatest contributions of your project?

The most significant outcome is that we created an excellent tool not only at Wake, but also across the state. It was also used a lot with state redesign in developmental math, which is the same course used at Wake. It's a valuable tool for students to monitor their math, and for those that want to enroll at Wake. They take the MOOC, and then they can test again after the remediation. Students often don't prepare very well for the placement test; we're marketing it to high school students.

SRI: What informed your design decisions for your courseware? What would you do differently to improve the overall design?

We loved the design and platform of the Udacity MOOC. The faculty worked really well with them to prepare the course to make sure it met their student learning outcomes. Wake did that in order to help a large student population that's at risk. When they thought open, they thought really open. The MOOC was initially thought of as a placement test, but it has become a supplement for other courses – accounting, business, and so on. Anyone can use it in any way they need to. One thing we'd do differently is to think more up front about the data we want, that's something Wake was naïve about. It would be nice to have known earlier the kind of information we'd want.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

Wake had been considering creating a MOOC in different areas, as more of a preparatory effort, not as a course replacement. Wake is more focused on getting students in their courses as quickly as possible. We lucked out and liked what we ended up with. How much work it will be is important. The MOOC was modularized and broken into pieces – it's important to have a good idea of how the MOOC will be used.

There was some interaction within the course. There was a discussion forum, and Wake monitors that loosely. Mostly the students ask each other questions and help each other along. The platform makes it easy to use. In this case, more instructor intervention wouldn't improve the course. If students are taking the MOOC as a review or as a supplement, then they have an instructor.

SRI: What student learning outcomes resulted from your project?

We don't have any granular data on student learning outcomes in the course because with Udacity it seems difficult to get to that level. I'm not sure if we would get granular data if we asked for it. Everything we've received so far has been report-based. We haven't asked for granular data, and really, there's no time to look at it. Wake would like to know how students who take the course do on the placement test. It's not possible to track back on the student survey to the placement course, which is disappointing because you can't see how it impacted student performance. Udacity doesn't track anything more that emails, but the survey does ask why they're taking the course. On the Survey Monkey data, out of 14,000 students, 3.5% of students reported that they were preparing for Wake Tech, 65% were taking it for personal knowledge, and 6-7% were completing whole course get a certificate from Udacity. If they wanted to take placement test for free, they would need a certificate. Otherwise it's $5 or $10 for the placement test.

SRI: What are your plans for sharing or furthering the work from this project?

We're excited about the data we're getting from Udacity. The average number of users average number of users increased from 1500 to 3000, a trend that's in line with the use in colleges during the semester. There is continued interest in the course. We received a phone call from a North Carolina testing site that wants to like our MOOC to their website. We did a lot of publicly with the Gates grant. We've had national attention on MOOCs and developmental MOOCs, as part of a Completion by Design grant – another form for getting the word out that we have a MOOC.

Interview with KARON KLIPPLE, Carnegie Foundation for the Advancement of Teaching

Pathways Initiative (Statway and Quantway Courses)

SRI: What do you consider to be the greatest contributions of your project?

We're opening the door for students to progress in their academics, career, and life who otherwise wouldn't have. When a student places into developmental math, there's about a 20% chance that that student will ever complete a college-level math course (even after 3 or more years). Now we're seeing 50% of developmental students earn college level credit in a single year…without this initiative they would be done. Opening the door to thousands more students and hopefully tens of thousands or hundreds of thousand more [is our greatest contribution].

SRI: What informed your design decisions for your courseware?

Our instructional design principles include starting with a realistic complex problem and allowing students to struggle productively with it. Giving students the opportunity to bring their own reasoning to bear on a problem allows them to create a conceptual framework for what they're about to learn. The Pathways use real, authentic contexts and real data to engage students in contexts that are inherently relevant and interesting. Within these contexts we layer support for basic mathematics concepts. We think every math topic should start with this kind of engagement and productive struggle.

It's a real shift for students and faculty. The teacher is much more in a facilitator's role. [Other design principles we use are] Explicit Connections—these help students see the connections within a lesson itself and how ideas in a lesson connect to what they learned before and what they're about to learn. And there's deliberate practice – we're strategic in giving students the opportunities to practice and test their learning, we give them fewer exercises than is typical but ones that push the boundaries in different directions from the base of what they learned in class. Many of these [principles] were informed by Jim Stigler's research with the TIMSS study. The use of this pedagogy and this changed role for faculty require significant professional development, which is an integral part of the Pathways.

Students use a Pathways workbook to support them in class. It guides them through productive struggle, group activities and other exercises. Online they have materials to support their in-class learning. [They have a] short encounter with a video, a reading or a set of exercises to prepare them for the next-day's class activities and then [there is a] follow-up through deliberate practice online. In addition, Statway provides a rich set of online resources, including readings, interactive checks for understanding, applets, and quizzes. This was built on the OLI Statistics course [for Statway]. In both Pathways [there is a] careful spiral to the curriculum so concepts are encountered again and again.

The Pathways are also different because we have embedded within them activities to support students' socio-emotional learning—around such things as their mindset about learning mathematics, their sense of belonging, and anxiety. This aspect of the Pathways is described well in the Silva and White paper.

That paper also describes how the Pathways is much more than a curriculum redesign. The Pathways is a collaborative effort of practitioners and researchers working together in a networked improvement community to solve the common problem of developmental mathematics success. Using the tools of improvement science we are able to take promising ideas—whether changes to curriculum or new instructional practices–test them in a small set of classrooms, and when appropriate, share them across our network of colleges.

SRI: What design recommendations would you give to technology developers for creating online learning environments or resources for first-generation college goers?

I almost think you should think about the faculty dashboard first and then build the student experience to provide that. The real power is the combined interaction of the student and teacher together. Faculty need a nimble way to understand what their students are struggling with so that they can adjust their instruction and provide targeted supports accordingly.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

Design of online supports differs between Statway and Quantway largely because they had different development processes. Quantway is a quantitative reasoning course, which is a relatively new course in mathematics departments, and as a result there were fewer existing resources from which we could build. The field has only recently converged on what a quantitative reasoning course is. Statway was designed with pretty extensive online resources outside of the classroom, [for use] in the absence of teammates and instructor. We felt there would be somewhat less need for that [online practice outside of class] for Quantway students. There were different expectations for the course, but they adhere to the same learning design principles. Statway takes students placing into elementary algebra and engages them in college-level statistics on Day 1. Quantway replaces elementary and intermediate algebra with a single developmental math course organized around quantitative reasoning [in semester 1] and then a college-level course in semester 2.

SRI: What student learning outcomes resulted from your project?

There is a report in the works from the analytics department around a common summative assessment given to all Pathways students. We worked with NIC [networked improvement community] practitioners to develop common assessment items that target key learning outcomes in each Pathway. We partnered with a team to edit [the items] for language and literacy. For psychometic analyses, we field tested [all items] with a comparison sample who had taken a college-level mathematics or statistics course. We found that Pathways students did just as well or better than the comparison sample across the items on each summative assessment.

A second report will compare the success of the Pathways students with similar students who took the traditional course sequence. By working together with our partner colleges we have access to a rich array of data. We have created a comparison population by doing student-by-student matching along 44 different characteristics. We are finding that students who take Pathways go on to earn more college-level credits than their comparison peers.

SRI: How will the work be sustained?

We think the Pathways have shown tremendous success, and our goal now is to make the Pathways available to more students.

If the colleges currently in our network were to make the Pathways the default for their developmental math students, we would reach 50,000 students. There are two colleges in the network that have already made Statway or Quantway the default course for students. We want to learn about how they were successful in making that happen. It requires buy-in from a variety of people with different motivations. We need to understand those motivations and how to meet them. If we don't scale to this extent, then the Pathways remain a project or a boutique option.

A second strategy is to look at how we bring new institutions into the network and to focus our attention on bringing in entire systems or states of colleges, say [through] legislative initiatives. If we do both of these, I think we'll reach 100,000 students a year.

SRI: What are your plans for sharing or furthering the work from this project?

Our goal is to help more students succeed in math and for more students to have access to the Pathways. We've just talked about ideas for scaling our network of colleges. Another idea is to bring the Pathways to more students through an online environment, and that's something we're just beginning to explore. We're trying to figure out, for example, how to provide the same opportunities for productive struggle and create the same sense of belonging online as students currently experience in class.

Interview with GARY LOPEZ, Monterey Institute for Technology and Education

NROC Project and DevMath

SRI: What do you consider to be the greatest contributions of your project?

The right business model is our greatest contribution. The value is in the process of prototyping, testing and then retesting, while all along building a community of early adopters and creating market buzz. We develop the strong involvement of our users to get constant feedback. Not just "Is it about right?" - but it gets down to the details of each product feature. Students completely space out if they hit a word they don't know. The argument from those that do traditional software development has been "You're wasting a lot of money, a lot of time." Our argument has been "This iterative development approach builds a product that is closer to what the users are requesting, while simultaneously accomplishing market development and creating a community of early adopters." When you combine software development and market development together in this manner, we would argue that you're actually building the software and the business faster and cheaper. For example, we haven't even released the final version of DevMath, and we have two states using it [statewide]. It's a different business model. It takes advantage of the social network and community that has become how the world works.

You put those two together, product development and market development, you get this new kind of educational publishing business model. I think this [combination] is going to be used in the future. While digital products have no unit costs as is the case of analog products, there is still a cost to design and develop educational materials that needs to be recovered, those don't go away. Somewhere along that value chain someone has to pay for it. Since only individuals and institutions have buying power one or the other (or both) must foot the bill for education products developed using this new business model. In keeping with the tenants of OER [open educational resources], we have chosen to keep our educational materials free for individual to use, but institutions must join NROC and help support the effort via a membership fee. On the positive side, the membership fee for access to all NROC assets is still a tiny fraction of what it costs to acquire similar educational materials from a commercial publisher.

We are not advocates of open resources, OER, as much as we are advocates of new business models. OER is the easiest low-hanging fruit on the way to a changed game for digital content on the Web. Open in the true sense of open where everyone changes the content is a problem for version control, especially for curricular materials where absolute compliance to standards is important. To allow customization or personalization, educational content needs to be open enough to be rearranged, but only within the confines of the curriculum.

SRI: What informed your design decisions for your courseware?

If you want a student to get engaged in education, make them care about it. The education experience needs to be something that resonates with that person. Educators and the publishers of educational materials haven't spent any time or money trying to figure out how to make education connect with students. That's where we as educational developers should be spending most of our time. The [math] concepts are really not that hard, but we make them much harder to understand by using a lexicon and exemplars that are foreign to most U.S. students. It is as though we are teaching a foreign language like Turkish.

The [DevMath] interface allows the student to have a goal tied to their aspiration, such as the school or program that seek to enter and see rapid progress toward that goal. The ability to iteratively go and raise your score has an element of gaming satisfaction. This is making progress through learning objectives. It seems to be engaging. DevMath has a very small dropout rate, unlike anything else we've worked on. DevMath uses video that is highly produced and really engaging at the beginning to explain the concept you're about to learn and exemplars of it. We use video for storytelling, and they're never more than 6 minutes. We don't use it for teaching; video doesn't do that well. We spend a lot of effort on getting the words, exemplars and cultural references right. And then there's the presenter—we went through 100 finding one with whom the students connected.

SRI: What suggestions would you give another organization that was

doing the same kind of work?

When we developed the NROC Algebra program, we were six months into it before we had the iterative feedback loop in place and that was a disaster. I wrote the first scripts and watched students nod off. They didn't know what I was talking about. I threw away six months of work and started over.

I would suggest building an iterative feedback process, including a definition of what the product is, from Day 1 with those who are going to be using it. You have to put a straw-man up that's informed by the subject and the technology resources available. That sets some boundaries but within them allow the users—administrators, teachers, and students—the opportunity to give feedback from Day 1. You can't be afraid of being called wrong. You have to be willing to start over, and be aware that every 30 days you may need to make adjustments.

You need a product bible with all of the key specs (e.g., editorial guidelines, design features, etc.) for guiding product development. It needs to be on the desk of the editorial staff, the design staff, the coding staff, etc., so that everybody has the same guiding vision. That bible gets refined by these iterative focus groups. If you do that, chances are you're going to publish something that's going to be "righter than wronger."

It's a lot of work. It kind of flies in the face of the way academics think…that scholastic precision is the only thing that matters. While academic/curricular content sets the informational/conceptual boundaries, it is only part of what makes good teaching materials and ignores or denies the importance of your ability to communicate and people's ability to learn. From the standpoint of the teachers, you've got to give them the tools that they know will work every day. From the standpoint of the students, you have to give them the material in a way they can understand it. And from the administrators' standpoint, you've got to give them the business case for really using it. Those three stakeholders should come to the table and be able to direct what you're doing. If one of them doesn't get what they need, you fail.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

Nationwide pilot studies identified four groups of students that gain benefit from DevMath. The use cases for each are different.

DevMath can be used in high school to ascertain if a student has sufficient math mastery to meet graduation requirements, and/or sufficient math mastery to eclipse the college math readiness cut score required by the higher education school or program the student seeks to enter. DevMath can be used in a traditional classroom as part of an emporium model in alternative schools or for online learning with a tutor.

Students who have been admitted to selective 4-year colleges but told their test scores suggest they may need math or English remediation can go to an DevMath site set up for that college to improve their readiness before getting [to the college] to avoid remediation. This is what they did in University of Montana. This works really well. At University of Montana 86% of DevMath users hit their target score and passed the math placement test. And 92% of the 35 students who achieved their Ed Ready goal score passed their first college math course at the University of Montana, a larger fraction than entering freshman at UM [University of Montana] who did not require remediation or use DevMath.

At community colleges and other nonselective colleges, some fraction of entering students may be able to benefit from DevMath on their own, but probably a smaller percentage than at selective colleges. Students at nonselective colleges are more likely to benefit from having a human teacher and a social situation. DevMath is more likely to be used in a boot camp or emporium situation with a good 3-month block. I'd suggest an emporium model with a teacher three times a week; that would probably take care of everyone in a semester.

The next step up would be to marry DevMath to the program of study—for example, teach welders what they need for their profession. They eventually need algebra, but only do that about a third of the way through their community college program. Do geometry later when they need it. You should never have to stop and do math as separate from your program. If it's in context, it makes sense to everybody.

We did have some successful pilots at ABE [Adult Basic Education] centers. Sometimes there are math instructors at these centers, but much more often an ABE student is working on their own—[doing] independent study. The challenge is to devise an approach to help keep the student on task and to answer any math questions that might arise.

SRI: What student learning outcomes resulted from your project?

[We] participated in study of math software for teaching in middle and high schools statewide in Utah. The governor funded the STEM Action Center from Utah State University look at various widely used products—ALEKS, Think Through Math, Math180, and DevMath. Ed Ready was used in grade 10. Volunteer schools used the math software products with students over a three-month period, fall semester 2013. DevMath did really well. Students that used DevMath progressed 15 times faster [in math achievement] than the current way math is taught in Utah.

SRI: Can you describe any unexpected outcomes, either positive or negative, that resulted from this work?

We saw in the University of Montana data that most of the entering freshman students, about 90%, spent something in the neighborhood of 5-50 hours working through their DevMath study plan with a median of 12 hours. We were surprised how short it was to avoid two semesters of remediation. Then there's the 5-10% who need much more time. They take hundreds of hours, but that's still a matter of weeks not semesters.

SRI: What are your plans for sharing or furthering the work from this project?

We've been under the radar for 10 years but now it's time to be seen. The State of Montana is rolling out DevMath statewide in the 2014-15 school year, as is the State of Utah. NROC won the College Readiness RFP that Utah put out last year. The Utah Education Network (UEN), the State of Utah DOE, and the STEM Action

Center (out of the Governor's office) teamed up to make DevMath available statewide to both secondary and post-secondary students. The states of Nevada, Maryland, and Tennessee are also considering statewide rollouts of DevMath, as is the California Community College System.

The Gates Foundation makes a lot of investments, and some of them work out and some don't. I think if you look at the grantees that receive the investments, typically they're small groups like ours, for whom building new things and learning new things comes easy, but there's no infrastructure for marketing those things. How can the successful products and services that arise from Gates Foundation investments get distributed nationally, to those that need it? The Gates Foundation could have a marketing/communications/distribution group dedicated to taking successful Gates-sponsored products and systems to market. In the VC [venture capital] world, a bunch of guys in a garage come up with something really cool. You give them money to do more but bring in a business manager, almost always a sales and marketing person. The Gates Foundation is acting like a VC. You don't give up after POC [proof-of-concept].

Interview with CATHY SANDEEN, American Council on Education

MOOCs for Credit Research

SRI: Can you describe the goals and objectives of your courseware project?

ACE launched our multi-faceted MOOCs for Credit research project to ask the question: What are some of the innovative methods for graduating more students and how might MOOCs contribute to this goal? We're looking at the credit approval process as a means for providing formal recommendation for credit, and applying that to a MOOC. The project also included developing a framework for evaluating learning pedagogy and interactivity within MOOCs. In addition, the project included convening a group called the "Presidential Innovation Lab." It's a collaborative effort between stakeholders to look at the academic potential of MOOCs. For the benefit of [academic] institutions, and to drive their agendas, institutional leadership needs to get involved – they have communication channels that are more influential. We know more about MOOCs and the role they might play in traditional degree programs now than we did at the beginning. The effort was a real "jump in feet-first" activity for everyone. We're all trying to understand MOOCs, OER [open educational resources], and technology-enhanced pedagogy, and it's rapidly evolving.

SRI: What are some of the current challenges with using MOOCs in higher education?

MOOCs have gone through some changes over the past year. The original MOOCs were more about openness and size of enrollment with a peer-distributed learning model. Later they focused on the video lecture approach for delivering content. Now MOOCs are becoming disaggregated—used in different ways at different institutions. For example, some colleges and universities import MOOCs developed at other institutions, or use some part of a MOOC as a supplementary resource in a flipped classroom format. It was just a matter of time before some of the students ask, “How can we get credit?” That led us to our research project. Some students are completing MOOCs as part of the admission process to show they can complete college-level work. There’s also the issue of “credit mobility.” It’s a major issue with enrollment patterns of the majority of students today. Most students will attend multiple institutions before graduating. We need to figure out a way to credibly allow students to transfer more credit. Faculty ultimately have authority over this process and they should. There are valid academic considerations in granting transfer credit. For example, it may not be a good idea to allow a student to transfer in credit from a technology-based course taken 12 years ago. That content is probably obsolete. But if a student completed one or two English composition courses, should they have to take another writing class?

There’s also the issue of intellectual property embedded in courses. Typically, faculty consider the course syllabi to be his or her IP. Institutions generally require faculty to assign rights to the IP to the institution if it is investing in an online course. If faculty leave, [the institution] doesn’t want their investment to “walk out the door.” In the case of MOOC platforms, Coursera and Udacity, the institution still owns the content, and anyone reusing it has to acquire a license for that. This arrangement can present challenges. The exception would be edX an open platform where the content and platform are an open resource.

SRI: Can you comment on some of the courseware design features or implementation strategies in the MOOCs you have reviewed?

MOOCs are designed to be used multiple times, though not forever, without ongoing faculty involvement. In most cases, courses will need to be revised and improved over time. Will the faculty member who originally created the MOOC be willing to put in the time for ongoing maintenance of the course? This is an open question. I guess it depends on the subject matter and how timeless the subject might be. Some courses can be repeated routinely with little change to the content. Some MOOCs are very specialized. Most MOOCs appeal to the “leisure learner” segment, individuals who have earned one or more degrees. There are some MOOCs that have been designed to be used broadly and may be better suited to undergraduate degree attainment. When we talk about general topics like statistics, that’s maybe aligned with a major textbook, there is generality and perhaps broader demand. Current research projects involving MOOCs are looking at efficacy, to see if students can learn this way. One of the happy surprises [of MOOCs] is the focus on the efficacy of teaching. For example, there’s a professor at MIT who taught a required chemistry course through edX. He’s been teaching this same lecture-based course for years, and then he tried it in a MOOC. He found that with the way the MOOC was designed to require mastery of each concept before progressing through the course, students were learning more in the MOOC-based model so he started implementing some MOOC techniques in his face-to-face course.

SRI: What are your thoughts on building knowledge for the field?

The findings and discoveries from these research projects are not shared for wider implementation. Grantees generally write a conference paper to fulfill a deliverable on the grant. That's great, so now what? We have thousands of "promising practices" and "islands of excellence" with everyone reinventing the wheel. It seems pretty inefficient. We need to look at research that's been done in a curated way, so people don't have to read hundreds of research papers. I would love to see something that aggregated and rationalized all this work. Not just MOOCs, but all research about technology interventions that improve persistence, success, and completion of postsecondary courses. Anecdotal evidence is valid, but having results distributed broadly and curated in some way is increasingly important. We have to move beyond the anecdotal and show some scale. I am also concerned about all the talk of the potential of "big data." Sure, the data are there, but knowing how to use it in a meaningful way is our major challenge moving forward.

Interview with PAUL STACEY, Creative Commons

Open Professionals Network (OPEN) - Raising the Floor with C3T. The OPEN consortium includes Creative Commons, Stanford University, Carnegie Mellon University, Center for Applied Special Technology, and the Washington State Board for Community & Technical Colleges.

SRI: What do you consider to be the greatest contribution of your project?

OPEN is supporting grantees participating in DoL's $2 billionTAACCT grant program. The DOL TAACCCT program is focused on moving people from unemployment or low-wage work to high-wage work in high growth industry sectors. The OPEN team is helping grantees with include requirements for use of the Creative Commons CC BY license, Open Educational Resources (OER), use of educational technology and online learning, data analytics, and using Universal Design for Learning and making curricula accessible. All of these are important contributions. From a Creative Commons perspective our big contribution has been to provide CC BY technical assistance to over 800 community colleges all across the US. All new resources colleges create with TAACCCT grant funds must be licensed CC BY as a requirement of the grant. CC has helped all these colleges learn about who CC is, what the CC BY license means, and how to apply CC BY to educational resources from a practical implementation perspective. CC has also provided assistance for these colleges on understanding that these CC BY licensed educational resources constitute Open Educational Resources and the big picture and benefits associated with that. One of the exciting things about this is that these programs and OER are in vocational areas where few OER currently exist. CC also provided technical assistance to DOL for defining requirements and selection of a repository and creation of a TAACCCT microsite where these resources will be housed. Collectively, so far, the colleges have already created 1100 programs of study (expect 2000 when all is done) with

80,000 enrolments, 27,000 of whom have already completed credentials leading to jobs in high growth industry sectors.

SRI: Can you describe any unexpected findings, either positive or negative, that resulted from this work?

One of the interesting findings is that TAACCCT grantees across many states are creating curricula for similar high growth industry sectors including health, advanced manufacturing, energy, transportation, and information technology. This means that multiple colleges across many states creating programs for these sectors are developing curricula in similar fields of study and that there is a lot of opportunity for sharing, coordinating development, and making use of Open Educational Resources from earlier rounds of the TAACCCT program and from the OER space in general. However the effort to network the TAACCCT grantees together and convene distributed groups for sharing and communication is big. In addition, development of programs and curricula has traditionally been done in an autonomous fashion. Shifting to reuse and a collaborative approach requires a break with autonomous practices, not-invented-here thinking, and more effort around planning and coordination.

Open Source Software [like Linux] has succeeded when it rallied a community of developers around continually working on and improving the code. We need a similar community building process for education where faculty in the same domain, but across multiple institutions, work collaboratively on developing curricula. This is starting to happen around open textbooks but needs to expand to full courses and programs.

When you think about the Bill & Melinda Gates Foundation's goal around scale, an open license is an enabler. It is far easier to scale something that is licensed to be shared, reused, remixed and redistributed than it is to scale something that is closed all rights reserved. To achieve scale [the Foundation] should adopt requirements that deliverables and research results produced with Foundation funding be openly licensed. This will generate greater distribution, use, and scale.

SRI: From your experience, how important is it to have a face-to-face component that supplements online learning?

It's critical that education involve social learning where students interact with each other and the instructor, not just the content. There are some unique aspects from being face-to-face, but online has it's own affordances. I find the question of whether online could be as good as face-to-face misleading as there's lots of terrible face-to-face education, not only in effectiveness but also in terms of being social. Its not face-to-face that is essential but social learning that is essential. Online learning and research about online learning has a long history of showing the benefits of social learning. These need to better leveraged than they are currently.

Another aspect to consider is encouraging colleges to do more to differentiate their online offering. There's not a distinctive university or college character for their online courses. I'm looking for institutions to create online courses that convey what's unique to that institution. If you look at the money invested in campus [physical] infrastructure, why not spend similar money to create a distinctly unique expression of the character and learning experiences of the institution online?

SRI: What design recommendations would you give to technology developers for creating online learning environments or resources for first-generation college goers?

We've been encouraging grantees to develop learning experiences in partnership with industry partners so their course examples are derived from the relevant industry. In some cases, the resulting curricula and program is something the industry partner as well as the college can use. Creating dual-purpose curricula is a good thing. Embedding examples from employment and jobs creates motivation. It helps to bring the industry partner to the table starting with the analysis and design of a course. It also helps to design for learner variability rather than a one size fits all. I'd encourage team work, social learning, and utilization of peer-to-peer expertise.

Interview with SASHA THACKABERRY, Cuyahoga Community College

Pre-Algebra MOOC

SRI: What do you consider to be the greatest contributions of your project?

The Pre-Algebra course turned out to be a great success. We really thought about our audience, the college students in the course. We ended up with a worldwide MOOC model, which is what we're seeing more of in college readiness for developmental education. The short time frame of the MOOC course contributed to its success; it was 4 weeks in length. Students who didn't complete in the 4 weeks could re-register and take the MOOC again the next 4 week session. Attendance really starts to dive after four weeks in traditional MOOCs, so we weren't interest in a 14-week course.

SRI: What informed your design decisions for your courseware? What would you do differently to improve the overall design?

The game mechanics were a key design feature; it was great for our audience. We started with an xMOOC self-paced, competency-based design and added on game mechanics. The game mechanics allowed us to create a safe-failure environment to improve student persistence and resilience.

In terms of improving design, we need to think about who our students are, and design for them. The use of game mechanics allowed us to leverage this student-centered perspective. This relates back to the affective components, such as grit, and how that impacts students who have struggled academically. We see this all the time anecdotally, but it would be good to collect data, either with a survey or an exit interview.

SRI: What guided your decisions for course implementation? Did you rely on a specific implementation model when implementing your courseware?

For our implementation model, ours is the only MOOC that has game features. It has competency-based game mechanics. You could not proceed to the next level until you got 80%. You work through the material, take a test, and if you don't finish it, then you take it again. The course was offered in four-week segments. The big iterative part of the course was how it was facilitated. The discussion board was monitored. Faculty created videos and sent out emails to connect with students and remind them to come back, and ended up developing more of those opportunities throughout the multi-session run.

The whole course also had to be accessible for it to run. There was little money attached to the project, so we spent a lot [of college funds] on resources. But without the support of the Gates Foundation and their belief in the model, we wouldn't have been able to do it. Courses need to be iterative – there are academic challenges we need to work through.

SRI: What student learning outcomes resulted from your project?

The hosting LMS [learning management system] had some data pull challenges. We got some data that didn't make any sense, and we kept getting the wrong CSV [comma-separated values] files. We wanted to connect an anonymous survey to student demographics and connect survey results with student achievement. We had to go back and get a unique identifier for students in order to link to students and demographics—it required considerable work to sort it out.

SRI: What are your plans for sharing or furthering the work from this project?

It depends on where the field is going. There's a lot of excitement about MOOCx, but there are so many people who aren't in the trenches. We need more people discussing the ground-level challenges. Without a voice on the ground, you can't really study this stuff.

I really think that we're going to see two kinds of MOOCs moving forward. The first is college readiness and development, and the second is higher ed MOOCs. We need to think about the audience. Funding needs to invest in institutions that have real challenges and are willing to try new things.

Interview with CANDACE THILLE, Carnegie Mellon University

Community College OLI

SRI: What do you consider to be the greatest contributions of your project?

A proof of concept that it's possible to move away from the model of individual faculty course development to a community-based research activity. We brought together multiple faculty from multiple institutions along with people with expertise about learning and software and articulated shared learning markers to design toward. [We showed we] Could get insight into how to support learners in a diversity of institutions.

SRI: Can you describe any unexpected findings, either positive or negative, that resulted from this work?

We found that counting students' OLI use more toward the course grade didn't drive more student use, but the degree to which faculty used the [OLI] Learning Dashboard did predict student use. What faculty write on their syllabus about what counts seems less a cue to students than what faculty mention repeatedly. Faculty who actually logged in with some frequency had students who used the course materials more and did better. It created coherence between what faculty were teaching and what's in the OLI course.

SRI: What were the main methodological challenges you encountered when evaluating the outcomes of the courseware you developed and/or implemented?

Intent-to-Treat [the fact that some faculty assigned to OLI don't really implement it] is probably the biggest challenge. The other challenge was the incredible variation within treatment condition and the lack of control classrooms. I wonder if controlled experimentation is the right approach. It may be too hard to pull off to get results that are usable. You could try and make inferences by introducing different learning approaches within the environment. A kind of dose/response analysis. I think the question of whether online is better or worse is not really that interesting. How do we design these environments to improve teaching and learning outcomes is the more interesting question. You just need everybody to have identifiable treatments.

SRI: Did you recommend a specific implementation model for your courseware?

I wanted to look at the effectiveness in the real world when you let faculty use the OLI course materials as they liked. I wanted to look at both effective faculty use and effective student use in a kind of 2 X 2. We didn't restrict faculty at all, so we ended up with an Intent-to-Treat study at the faculty level, which mediated what happened at the student level.

I wanted faculty to look at the knowledge state first, then see those who are struggling and click in to see if it's lack of effort or the online activities just don't work for them and then try to get insights into where students are struggling. Few faculty do this. Most faculty click in to see how many activities the students did.

Having faculty grade on how many things you did is not the teaching and learning model I'm trying to support. I think you need faculty development sessions on the use of dashboards. We're asking the faculty to learn a new practice. Just telling someone to do that is not sufficient. We need to look at faculty as though they are the students for this change. How do you scaffold this behavior? We need to start with the faculty where they are. We don't want to support just what they currently do without a way to support the transition to where they need to go.

We recommend [a specific model] especially for first generation or community college students who have been trained with a particular learning model (your job is to show up, listen to what I tell you, write it down, and write it again in about a week). A lot of students think that's their role. For faculty to say, "Actually you want facility with this tool [statistics] that helps you apply it to things in the world that are meaningful to you. To have that facility, this is how we'll know when you're there. Our job is to move you from wherever you are to that level of facility. We share responsibility for getting you there. We have multiple resources to help you get there. One of them is this online course. One of them is your peers. We want to get the whole class there. Another resource is me." The instructor shows the learning outcomes for the first module and sends students to work online before coming to class the next week. "Do your best to turn your [OLI] dashboard green. We'll have a much richer conversation when you come to class next week if you're green. When class starts we'll start off by assessing where the class is." That's a behavior change for both faculty and students. When faculty first start using OLI courses, their teaching evaluations go down. Students have the idea that the faculty's job is to teach them. Faculty successful with OLI spend considerable time throughout the course reminding students of this [new] role. They tell them it will feel hard.

SRI: Are these recommendations specific to certain subjects?

I think it's less about the [subject] domain than about the type of knowledge that's the target for instruction. Humanities courses have higher frequency of a certain type of knowledge than a science course might. We need to look below the course subject level to the type of knowledge and learning outcomes. Being able to articulate what you want students to be able to do and know is a design process that should work in any subject area.

SRI: What suggestions would you give to a colleague or another organization that was doing the same work?

I think there's a tension between creating well-designed full courses that demonstrate learning and creating really good authoring tools so people can create stuff and offering a junk drawer of learning activities. One has the problem of adoption, and the other has the problem of inappropriate adaptation. I also think talking about faculty as if they're a monolithic group is fallacious. There are some faculty who have a clear idea and want to be able to modify [an online course]. There are faculty who just want help teaching their assigned course. Just give me something that works. We need an environment that supports faculty all along that continuum. When we give faculty the tools to adapt, it [the adaptation] needs to be driven by an evidence-based practice approach. The tool needs to collect evidence on whether that change had positive of negative impact. It needs to be part of the tool. There is not a silver bullet here.

SRI: Can you describe any findings with respect to supporting faculty in transitioning to this new way of teaching?

Faculty could support each other in that change. Faculty say they want to do that, but no one has time. We tried to build that into the environment. We put a discussion forum for faculty into the OLI environment. Faculty were asking for it. It appears on the instructor home page. When you create an instructor account, you get joined to a community for OLI overall and for the course you're teaching. We had some seeded questions and posted syllabi from experienced OLI faculty and still faculty didn't persist in forum use past the first week of class. You can't just build a community by creating a place for it.

Faculty who were involved in the course development did support each other . . . . In OLI the barrier to entry was too high. You had to be part of the OLI development team at some level [to really be part of the faculty community around the OLI courses]. [We] Didn't build authoring tools for mere mortals to use to allow them to replace units of the course. OLI does allow selecting and sequencing modules but not adding or modifying something. Open edX will allow that.

Interview with CAROL TWIGG, NCAT/Changing The Equation

Developmental Math Course Redesign

SRI: What do you consider to be the greatest contributions of your project?

NCAT has shown how [colleges] can achieve better learning outcomes while reducing costs in developmental math. The fact that we have done this is the main achievement.

SRI: Can you describe any unexpected outcomes, either positive or negative, that resulted from this work?

The Change article on Changing the Equation discusses 10 things we learned. We were surprised by the level of difficulty community college faculty had dealing with data, whether assessment or cost data. We've collected the same data from hundreds of colleges. They [community college faculty] really struggled with it… they had no sense of how to use data to make improvements. That's a pretty big problem…this inability to use data to shape what you're doing and measure the outcome, that was a surprise to us.

SRI: How important is it to have an external organization organizing and supporting the change process?

It is totally critical. Most colleges and universities are not systemic change agents. They have too many things going on in their lives. It's very difficult to focus on a radically different change project; they just don't do it. We encourage institutions to form teams and include academic administrators, IT staff as well as faculty, which they do, but academic administrators tend to [just] pay lip service to the project. They don't keep an eye on it. So when issues like faculty resistance come up, the teams tend to drift off and not finish the project unless someone—in our case, an external organization—monitors the implementation.

SRI: Did you specify a specific implementation model for the hybrid courses?

Yes, what we call the emporium model. There are several key modifications [to the way mathematics is usually taught in colleges]: mandatory attendance in a computer lab or classroom at least 3 hours weekly (freshmen do not do “optional”); modularized curriculum with individualized pacing; and mastery learning. We have found that redesign will work in any discipline and have done hundreds of these. We alert institutions to critical implementation issues, which apply across the board at research universities and community colleges. The two most critical conditions [for successful implementation] are leadership and consensus.

SRI: From your experience, how important is it to have a face-to-face component that supplements online learning?

We believe the face-to-face component is critical. We don’t think teaching developmental math in a fully online environment will work for most students. Human contact, keeping them on point and encouraging them, is critical.

SRI: What design recommendations would you give to technology developers for creating online learning resources for use in blended college courses?

Mastery learning is not as important in subjects other than skills-based courses. We think it’s valuable, and we suggest it, but we don’t require it. We think it’s critical in skills-based courses like developmental math and developmental English. But in general, people don’t use it in other subject areas.

SRI: Can you comment on how big a change it is for the IHEs you work with to track student learning and student course completion rates for different iterations of their courses?

It’s a major change. Most don’t even look at pass rates. A few years ago when starting these programs, we had a meeting with 12 chancellors of university systems and asked them to report the pass rates of their freshman math courses… They were absolutely stunned at how bad their pass rates were. If you ask the average college administrator, “What are your pass rates for freshman math or developmental math?” they will not have a clue. To get them to look at what students are actually learning and why they’re learning or not is a major change.

SRI: What are your plans for sharing or furthering the work from this project?

Over the past 14 years, we have done hundreds of redesigns. We’ve decided to shift our emphasis from conducting redesign programs to a focus on change strategies and the creation of resources that people can use without us. We’ve created “cookbooks”—how-to guides, two for math and one for other academic areas. We have appointed about 50 NCAT Redesign Scholars that institutions can hire as consultants. We don’t have as many examples as we would like in developmental English, so we’re considering doing a redesign program in that area with additional institutions.

Appendix B: Meta-analysis Data Tables

Exhibit B-1. Postsecondary Success Courseware Context of Use

Dimension	Courseware Context	Frequency*	Percent
Field of Use	Community College	85	61
	4-Yr College	54	39
Courseware Role	Whole-course model	16	12
	Course redesign	61	44
	Portion of course	18	13
	Supplemental resource	17	12
	Supports for course redesign	15	11
	Learning Analytics/Early alert	6	4
	Supports for peer learning	6	4
Academic Preparation	Low	108	78
	Mid	25	18
	High	6	4
Platform/ Technology	MOOC	15	11
	OLI	9	6
	Adaptive technologies	60	43
	LMS	55	40
Subject Area	Mathematics	79	57
	Statistics	2	1
	Computer Science	4	3
	Physical & Life Sciences	22	16
	Social Science & Business	12	9
	English & Humanities	16	12
	Mixed	4	3

**Includes all courses across grantees (N = 139)*

Exhibit B-2. Postsecondary Success Instructional Design and Technology Features

Dimension	Courseware Feature	Frequency*	Percent
Pacing	Self-paced (open entry/open exit)	88	63
	Class-paced	30	22
	Class paced with some self-paced elements	21	15
Pedagogy	Expository	32	23
	Practice environment	93	67
	Exploratory	8	6
	Collaborative	6	4
Source of Feedback	Automated	126	91
	Teacher	2	1
	Peers	11	8
Communication Synchrony	Asynchronous & Synchronous	6	4
	Asynchronous only	130	94
	Synchronous only	3	2
Individual Learning Path	Mastery-based	59	42
	Learner choice	17	12
	None	63	45
Virtual Experience	Game feature	3	2
	Teachable Agent	8	6
	None	128	92

**Includes all courses across grantees (N = 139)*

Exhibit B-3. Postsecondary Success Courseware Implementation Practices

Dimension	Implementation Strategy	Frequency*	Percent
Modality	Fully online	21	15%
	Blended with over 50% online but at least 25% FTF	67	48
	Blended with 25-50% online	51	37
Student: Instructor Ratio	≤ 35 to 1	7	5
	36-99 to 1	30	22
	100-299 to 1	50	36
	300-599 to 1	21	15
	600-999 to 1	8	6
	≥ 1000 to 1	23	17
Instructor Role Online	Active presence online	0	0
	Small presence online	15	11
	No presence online	124	89
Dominant Student Role	Listen/read	31	22
	Complete problems/answer questions	95	68
	Explore simulation/resources	5	4
	Collaborate with peers	8	6

**Includes all courses across grantees (N = 139)*

Exhibit B-4. Postsecondary Success Courseware Impacts on Course Completion Rate

Project	Course/Campus	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
ALMAP	ASU	0.310	0.120	0.075	0.545	2.583**	738
ALMAP	Essex	0.130	0.140	-0.144	0.404	0.929	408
ALMAP	Rio Salado	0.949	0.702	-0.427	2.326	1.352	34
ALMAP	Savanna Tech	-0.020	0.170	-0.352	0.312	-0.118	253
ALMAP	St. Petersburg ALEKS	-0.150	0.130	-0.404	0.105	-1.154	199
ALMAP	St. Petersburg LearnSmart	0.548	0.249	0.060	1.036	2.200*	98
ALMAP	UC Davis	0.190	0.180	-0.163	0.542	1.056	578
ALMAP	ALMAP AVERAGE	0.134	0.068	0.001	0.267	1.974*	2,308
NROC	Daemen College	0.094	0.290	-0.475	0.664	0.325	28
NROC	Jackson State University	0.289	0.201	-0.104	0.682	1.443	74
NROC	NROC AVERAGE	0.226	0.165	-0.097	0.550	1.372	102
ITHAKA	UMD Blended MOOC	0.155	0.078	0.002	0.307	1.981*	786
ITHAKA	UMD Blended MOOC AVERAGE	0.155	0.078	0.002	0.307	1.981*	786
NCAT	Bowling Green Math 55 fall 2011	-0.192	0.132	-0.451	0.067	-1.453	138
NCAT	Bowling Green Math 55 spr 2012	-0.332	0.181	-0.687	0.023	-1.836	123
NCAT	Bowling Green Math 65 fall 2011	0.313	0.121	0.075	0.551	2.578**	166
NCAT	Bowling Green Math 65 spr 2012	-0.014	0.174	-0.356	0.328	-0.080	82
NCAT	Cossatot Essential Math	-0.515	0.271	-1.047	0.016	-1.900	30
NCAT	Cossatot Intermediate Algebra	-0.588	0.204	-0.987	-0.189	-2.886**	55
NCAT	Cossatot Introductory Algebra	-0.197	0.246	-0.680	0.285	-0.803	56
NCAT	Guilford Essential Math	-0.403	0.057	-0.514	-0.291	-7.098***	641
NCAT	Guilford Intermediate Algebra	-0.517	0.117	-0.746	-0.288	-4.431***	137
NCAT	Guilford Introductory Algebra	0.611	0.067	0.479	0.742	9.108***	409
NCAT	Laramie Math 900 fall 2011	-0.401	0.150	-0.695	-0.108	-2.678**	114
NCAT	Laramie Math 900 spr 2012	0.022	0.170	-0.312	0.356	0.130	103
NCAT	Laramie Math 920 fall 2011	0.159	0.126	-0.087	0.405	1.264	180
NCAT	Laramie Math 920 spr 2012	-0.044	0.122	-0.283	0.195	-0.363	185
NCAT	Laramie Math 930 fall 2011	-0.516	0.157	-0.824	-0.208	-3.283***	102
NCAT	Laramie Math 930 spr 2012	0.044	0.186	-0.321	0.409	0.236	101
NCAT	LBWCC Basic Math	-0.515	0.184	-0.875	-0.155	-2.806**	61
NCAT	LBWCC Elementary Algebra	-0.669	0.174	-1.011	-0.328	-3.845***	81
NCAT	LBWCC Intermediate Algebra	-0.093	0.189	-0.465	0.278	-0.493	49
NCAT	Manchester Math 075	-0.155	0.099	-0.350	0.040	-1.556	245
NCAT	Manchester Math 095	0.047	0.068	-0.086	0.179	0.693	583
NCAT	Oakton Math 060	-0.400	0.134	-0.663	-0.137	-2.984**	142
NCAT	Oakton Math 070	-0.222	0.103	-0.425	-0.020	-2.149*	251
NCAT	Oakton Math 080	0.081	0.207	-0.324	0.486	0.391	60
NCAT	Oakton Math 110	0.092	0.103	-0.110	0.293	0.892	281
NCAT	Pearl River Math Fundamentals	-0.025	0.139	-0.297	0.247	-0.178	128

Exhibit B-4. Postsecondary Success Courseware Impacts on Course Completion Rate Implementation Practices, Continued

Project	Course/Campus	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
NCAT	Pearl River Beginning Algebra	0.083	0.090	-0.095	0.260	0.915	349
NCAT	Pearl River Intermediate Algebra	-0.082	0.062	-0.203	0.039	-1.334	488
NCAT	Robeson Essential Math	-0.288	0.099	-0.482	-0.094	-2.916**	220
NCAT	Robeson Introductory Algebra	-0.202	0.114	-0.426	0.022	-1.769	156
NCAT	Somerset Basic Algebra	-0.266	0.082	-0.426	-0.105	-3.245***	321
NCAT	Somerset Pre-Algebra	-0.311	0.078	-0.464	-0.157	-3.964***	357
NCAT	WVUP Math 011 fall 2011	-0.347	0.100	-0.544	-0.150	-3.451***	212
NCAT	WVUP Math 011 spr 2012	-0.404	0.148	-0.694	-0.115	-2.735**	83
NCAT	WVUP Math 021 fall 2011	-0.228	0.118	-0.459	0.002	-1.939	169
NCAT	WVUP Math 021 spr 2012	-0.531	0.200	-0.924	-0.139	-2.653**	64
NCAT	NCAT AVERAGE	-0.171	0.044	-0.256	-0.085	-3.916***	6,922
NGLC	Abilene Christian ACU	-0.110	0.174	-0.451	0.232	-0.630	136
NGLC	Abilene Christian CalU	-0.027	0.461	-0.930	0.877	-0.058	45
NGLC	Bryn Mawr	0.080	0.139	-0.193	0.353	0.571	270
NGLC	Carnegie Learning	-0.038	0.100	-0.234	0.158	-0.376	226
NGLC	Central Piedmont	-0.195	0.017	-0.227	-0.162	-11.786***	6,164
NGLC	Cerritos Beginning Algebra	0.507	0.206	0.103	0.911	2.459*	225
NGLC	Cerritos Business Fundamentals	0.112	0.064	-0.013	0.236	1.757	960
NGLC	Cerritos Business Management	0.013	0.127	-0.236	0.262	0.104	321
NGLC	Cerritos Dev Reading	0.075	0.230	-0.375	0.525	0.327	126
NGLC	Cerritos Dev Writing	-0.141	0.157	-0.448	0.167	-0.896	420
NGLC	Cerritos Intermediate Algebra	0.603	0.190	0.230	0.976	3.168**	295
NGLC	Cerritos Intro to Psychology	0.481	0.126	0.233	0.728	3.804***	226
NGLC	Cerritos Physical Geography	-0.241	0.102	-0.441	-0.041	-2.360*	466
NGLC	Chattanooga Fall 2011	0.197	0.035	0.129	0.266	5.663***	2,048
NGLC	Chattanooga Fall 2012	0.830	0.038	0.756	0.905	21.949***	2,036
NGLC	Chattanooga Spring 2012	-0.046	0.042	-0.128	0.035	-1.108	1,468
NGLC	Chicago Community College	0.486	0.049	0.390	0.583	9.847***	648
NGLC	Indiana FIU	-0.227	0.190	-0.600	0.146	-1.191	79
NGLC	Indiana IUPUI	-0.446	0.212	-0.862	-0.030	-2.101*	68
NGLC	Iowa CCOC English 105	-0.073	0.092	-0.253	0.106	-0.800	223
NGLC	Iowa CCOC Math 110	-0.095	0.134	-0.357	0.167	-0.710	154
NGLC	Iowa CCOC Math 121	-0.201	0.129	-0.453	0.051	-1.562	108
NGLC	Marist College	0.045	0.050	-0.052	0.143	0.909	1,547
NGLC	Missouri Basic English	-0.265	0.125	-0.510	-0.019	-2.110*	141
NGLC	Missouri College Algebra SEMO	-0.176	0.065	-0.303	-0.049	-2.711**	830
NGLC	Missouri College Algebra UMKC	-0.190	0.135	-0.453	0.074	-1.411	310
NGLC	Missouri Concept of Biology	0.157	0.105	-0.048	0.362	1.500	432
NGLC	Missouri General Chemistry	-0.035	0.068	-0.169	0.099	-0.512	751

Exhibit B-4. Postsecondary Success Courseware Impacts on Course Completion Rate Implementation Practices, Continued

Project	Course/Campus	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
NGLC	Missouri Human Anatomy	0.090	0.112	-0.128	0.309	0.810	163
NGLC	Missouri Information Systems	-0.130	0.143	-0.410	0.150	-0.909	316
NGLC	Missouri Intermediate Algebra	0.390	0.085	0.223	0.557	4.583***	497
NGLC	Missouri Intro to Psychology	-0.015	0.082	-0.175	0.145	-0.181	1,340
NGLC	Missouri Intro to Business	-0.215	0.154	-0.516	0.087	-1.396	65
NGLC	Missouri Health and Fitness	-0.166	0.161	-0.482	0.150	-1.030	600
NGLC	Missouri Oral Communication	-0.316	0.094	-0.500	-0.133	-3.376***	394
NGLC	Missouri Management	-0.540	0.160	-0.853	-0.227	-3.379***	171
NGLC	UW-Milwaukee Salon Biology	-0.203	0.124	-0.447	0.040	-1.636	111
NGLC	UW-Milwaukee Salon Chemistry	0.056	0.128	-0.196	0.307	0.435	178
NGLC	UW-Milwaukee UPACE UWM	1.388	0.192	1.012	1.765	7.224***	106
NGLC	UW-Milwaukee UPACE UNF	0.297	0.163	-0.024	0.617	1.816	78
NGLC	UW-Milwaukee UPACE UPR	0.195	0.222	-0.240	0.629	0.879	46
NGLC	Wake Forest	-0.036	0.129	-0.289	0.217	-0.282	403
NGLC	WICHE Fall	0.000	0.071	-0.139	0.139	0.000	473
NGLC	WICHE Spring and Summer	0.023	0.082	-0.137	0.183	0.280	348
NGLC	NGLC AVERAGE	0.055	0.072	-0.086	0.195	0.762	26,012
Pathways	Quantway 2012-13	0.766	0.041	0.686	0.845	18.854***	1,481
Pathways	Quantway 2011-12	0.832	0.054	0.725	0.939	15.286***	573
Pathways	Statway 2010-11	0.805	0.165	0.482	1.128	4.879***	928
Pathways	Statway 2011-12	0.943	0.138	0.672	1.213	6.833***	771
Pathways	PATHWAYS AVERAGE	0.798	0.031	0.737	0.859	25.642***	3,753
Overall	POSTSECONDARY SUCCESS AVERAGE	0.372	0.021	0.330	0.414	17.328***	39,883

Exhibit B-5. Postsecondary Success Courseware Impacts on Learning Outcomes

Project	Campus/ Course	Outcome	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
ALMAP	ASU	Grade	0.160	0.080	0.003	0.317	2.000*	738
ALMAP	Essex	Grade	-0.030	0.110	-0.245	0.186	-0.273	379
ALMAP	Rio Salado	Grade	0.633	0.722	-0.782	2.048	0.877	34
ALMAP	Savanna Tech	Grade	0.408	0.289	-0.158	0.975	1.414	152
ALMAP	St. Petersburg ALEKS	Grade	-0.050	0.100	-0.245	0.146	-0.500	165
ALMAP	St. Petersburg LearnSmart	Grade	0.070	0.129	-0.184	0.323	0.538	94
ALMAP	UC Davis	Grade	0.010	0.070	-0.127	0.147	0.143	538
ALMAP	ALMAP AVERAGE	Grade	0.049	0.043	-0.035	0.134	1.145	2,100
CC-OLI	CC-OLI	Post Assessment	0.070	0.060	-0.047	0.187	1.167	84
CC-OLI	CC-OLI AVERAGE	Post Assessment	0.070	0.060	-0.047	0.187	1.167	84
ITHAKA	UMaryland Blended MOOC	Grade	-0.018	0.050	-0.117	0.081	-0.355	786
ITHAKA	BLENDED MOOC AVERAGE	Grade	-0.018	0.050	-0.117	0.081	-0.355	786
NCAT	Bowling Grn Mat 55 fall 2011	Post Assessment	0.464	0.120	0.229	0.700	3.864***	138
NCAT	Bowling Grn Mat 55 spring 2012	Post Assessment	0.516	0.165	0.193	0.839	3.129**	123
NCAT	Bowling Grn Mat 65 fall 2011	Post Assessment	0.207	0.107	-0.002	0.417	1.939	166
NCAT	Guilford Ess Math	Final Exam	1.218	0.134	0.955	1.481	9.076***	641
NCAT	Guilford Intermediate Alg	Final Exam	1.004	0.177	0.657	1.351	5.668***	137
NCAT	Guilford Intro Alg	Final Exam	0.609	0.104	0.405	0.813	5.862***	409
NCAT	Laramie Math 900 fall 2011	Post Assessment	1.757	0.157	1.449	2.066	11.171***	114
NCAT	Laramie Math 900 spring 2012	Post Assessment	1.639	0.177	1.292	1.985	9.262***	103
NCAT	Laramie Math 920 fall 2011	Post Assessment	1.766	0.132	1.507	2.025	13.368***	180
NCAT	Laramie Math 920 spring 2012	Post Assessment	2.716	0.154	2.415	3.017	17.687***	180
NCAT	Laramie Math 930 fall 2011	Post Assessment	2.642	0.190	2.270	3.014	13.915***	102
NCAT	Laramie Math 930 spring 2012	Post Assessment	2.887	0.236	2.425	3.349	12.249***	101
NCAT	LBWCC Basic Math	Post Assessment	-0.326	0.163	-0.646	-0.005	-1.993*	61
NCAT	LBWCC Elementary Algebra	Post Assessment	-0.289	0.145	-0.574	-0.005	-1.997*	81
NCAT	LBWCC Intemediate Algebra	Post Assessment	-0.316	0.167	-0.643	0.010	-1.899	49

Exhibit B-5. Postsecondary Success Courseware Impacts on Learning Outcomes, Continued

Project	Campus/ Course	Outcome	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
NCAT	Oakton MAT 060	Post Assessment	1.230	0.131	0.975	1.486	9.426***	142
NCAT	Oakton MAT 070	Post Assessment	1.019	0.099	0.825	1.212	10.307***	251
NCAT	Oakton MAT 080	Post Assessment	0.873	0.176	0.528	1.219	4.954***	60
NCAT	Oakton MAT 110	Post Assessment	1.130	0.098	0.937	1.323	11.491***	281
NCAT	Pearl River Fundamentals of Math	Post Assessment	1.436	0.159	1.125	1.747	9.049***	74
NCAT	Pearl Riverer Beginning Algebra	Post Assessment	1.062	0.086	0.894	1.230	12.404***	349
NCAT	Pearl Riverer Intermediate Algebra	Post Assessment	0.622	0.067	0.491	0.753	9.301***	488
NCAT	Robeson Ess Math	Post Assessment	1.143	0.096	0.956	1.331	11.943***	220
NCAT	Robeson Intermediate Algebra	Post Assessment	0.144	0.280	-0.406	0.693	0.512	40
NCAT	Robeson intro Algebra	Post Assessment	0.583	0.104	0.379	0.787	5.603***	156
NCAT	Somerset Basic-Algebra	Post Assessment	0.610	0.075	0.463	0.758	8.102***	321
NCAT	Somerset Pre-Algebra	Post Assessment	0.744	0.073	0.601	0.886	10.233***	357
NCAT	Stark State MTH100	Post Assessment	2.371	0.069	2.236	2.505	34.516***	833
NCAT	Stark State MTH101	Post Assessment	1.592	0.089	1.417	1.767	17.843***	188
NCAT	Stark State MTH123	Post Assessment	2.548	0.214	2.129	2.968	11.899***	91
NCAT	WVUP Math 011 fall 2011	Post Assessment	0.825	0.094	0.641	1.008	8.802***	212
NCAT	WVUP Math 011 Spring 2012	Post Assessment	0.947	0.129	0.694	1.199	7.349***	83
NCAT	WVUP Math 021 fall 2011	Post Assessment	-0.039	0.104	-0.243	0.166	-0.371	169
NCAT	WVUP Math 021 Spring 2012	Post Assessment	0.338	0.144	0.056	0.620	2.349*	64
NCAT	NCAT AVERAGE	Post Assessment	0.891	0.222	0.456	1.326	4.013***	6,964

Exhibit B-5. Postsecondary Success Courseware Impacts on Learning Outcomes, Continued

Project	Campus/ Course	Outcome	Effect Size	Standard Error	Lower Limit	Upper Limit	Z-Value	Sample Size
NGLC	Abilene Christian	Post Assessment	-0.060	0.211	-0.474	0.355	-0.283	43
NGLC	Bryn Mawr	Grade	0.053	0.065	-0.074	0.180	0.817	270
NGLC	Carnegie Learning	Grade	-0.008	0.090	-0.186	0.169	-0.094	226
NGLC	CSU-N	Grade	0.491	0.028	0.436	0.546	17.469***	4,479
NGLC	Indiana	Post Assessment	-0.003	0.195	-0.384	0.379	-0.015	52
NGLC	Marist	Grade	0.136	0.044	0.051	0.222	3.130**	1,547
NGLC	Missouri College Algebra	Post Assessment	0.360	0.060	0.243	0.478	6.006***	575
NGLC	Missouri Chemistry	Post Assessment	0.586	0.055	0.479	0.693	10.736***	672
NGLC	Missouri Info Systems	Post Assessment	0.558	0.092	0.377	0.739	6.045***	316
NGLC	OhioLINK Fall 2011	Grade	-0.114	0.972	1.547	-0.114	0.972	24
NGLC	OhioLINK Fall 2012	Grade	-0.087	0.181	0.687	-0.087	0.181	233
NGLC	Open U	Grade	0.268	0.135	0.004	0.532	1.989*	86
NGLC	SUNY	Grade	-0.093	0.121	-0.329	0.144	-0.770	119
NGLC	U Mass	Grade	-0.209	0.234	-0.668	0.249	-0.895	30
NGLC	U Michigan	BTE analysis	0.092	0.031	0.031	0.152	2.964**	1,383
NGLC	UW-Milwaukee Salon Biology	Assessment	-0.165	0.075	-0.312	-0.019	-2.213*	312
NGLC	UW-Milwaukee Salon Chemistry	Assessment	0.003	0.103	-0.199	0.205	0.030	178
NGLC	NGLC AVERAGE	Various	0.110	0.072	-0.032	0.251	1.520	10,545
Pathways	Statway 2012-13	Final exam	1.244	0.042	1.162	1.326	29.711***	771
Pathways	Statway 2010-11	Accumulated college credits	0.670	0.037	0.598	0.741	18.310***	928
Pathways	Pathways AVERAGE	Various	0.918	0.028	0.864	0.972	33.336***	1,699
Overall	POSTSECONDARY SUCCESS AVERAGE	Various	0.470	0.019	0.432	0.507	24.603***	22,178

Exhibit B-6. Learning Outcomes by Courseware Role

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Course redesign/ Whole course	42	0.960	0.090	0.783	1.138	10.624***	100.610***
Portion of course/ Supplemental resource	11	0.024	0.026	-0.027	0.075	0.916	
Supports for course redesign/ Learner analytics/ Supports for peer learning	9	0.040	0.036	-0.031	0.111	1.105	

**p < .05; ** p < .01; ***p < .001*

Exhibit B-7. Learning Outcomes by Field of Use

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
4-year college	22	0.248	0.059	0.133	0.363	4.235***	28.218***
Community college	38	0.933	0.115	0.708	1.158	8.121***	

**p* < .05; ** *p* < .01; ****p* < .001*

Exhibit B-8. Learning Outcomes by Students' Prior Achievement

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
High/ Medium/ Mixed	18	0.557	0.124	0.315	0.800	4.508***	0.888
Low	44	0.699	0.086	0.531	0.868	8.139***	

**p* < .05; ** *p* < .01; ****p* < .001*

Exhibit B-9. Learning Outcomes by Technology Platform Type

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Adaptive	48	0.794	0.103	0.592	0.995	7.718***	13.378***
Other	14	0.225	0.117	-0.003	0.454	1.933	

*$p < .05$; ** $p < .01$; ***$p < .001$

Exhibit B-10. Learning Outcomes by Subject Area

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Mathematics	48	0.816	0.090	0.639	0.993	9.043***	41.927***
Other	14	0.116	0.059	0.000	0.233	1.958*	

*$p < .05$; ** $p < .01$; ***$p < .001$

Exhibit B-11. Learning Outcomes by Courseware Pacing

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Class-paced/ Mixed pacing	15	0.240	0.115	0.014	0.465	2.085*	12.867***
Self-paced	47	0.794	0.103	0.591	0.996	7.693***	

*$p < .05$; ** $p < .01$; ***$p < .001$

Exhibit B-12. Learning Outcomes by Pedagogy

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Practice environment	52	0.787	0.084	0.622	0.951	9.361***	71.695***
Collaborative/ Exploratory	6	0.006	0.050	-0.091	0.103	0.121	
Expository	4	0.041	0.043	-0.043	0.132	0.124	

**p < .05; ** p < .01; ***p < .001*

Exhibit B-13. Learning Outcomes by Individual Learning Path

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Mastery-based	37	0.999	0.114	0.776	1.222	8.777***	62.34***
Learner choice	3	-0.056	0.070	-0.194	0.082	-0.798	
None	22	0.186	0.089	0.011	0.361	2.081*	

**p < .05; ** p < .01; ***p < .001*

Exhibit B-14. Learning Outcomes by Role of Online Assessments

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Determine if student is ready for new content	45	0.857	0.097	0.668	1.047	8.860***	22.039***
Other	17	0.155	0.114	-0.068	0.379	1.363	

**p < .05; ** p < .01; ***p < .001*

Exhibit B-15. Learning Outcomes by Modality

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Blended with 25-50% online	14	0.048	0.071	-0.090	0.187	0.683	56.012***
Blended with over 50% but less than 76% online	44	0.895	0.097	-0.035	1.084	9.260***	
Fully online	4	0.125	0.068	-0.009	0.259	1.829	

**p < .05; ** p < .01; ***p < .001*

Exhibit B-16. Learning Outcomes by Student: Instructor Ratio

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
<100 to 1	13	0.357	0.196	-0.028	0.742	1.818	8.353*
100-299 to 1	26	0.941	0.153	0.641	1.240	6.160***	
300-999 to 1	14	0.622	0.177	0.276	0.968	3.521***	
>1000 to 1	8	0.397	0.150	0.103	0.691	2.650**	

*p < .05; ** p < .01; ***p < .001

Exhibit B-17. Learning Outcomes by Dominant Student Role Online

Group	Number contrasts	Effect size	Standard error	Lower limit	Upper limit	Z-value	Q-statistic
Complete problems/ Answer questions	50	0.816	0.086	0.648	0.985	9.518***	75.764***
Listen/ Read	6	0.062	0.023	0.018	0.106	2.746**	
Collaborate with peers/ Explore simulation or resources	6	-0.012	0.063	-0.134	0.111	-0.190	

*p < .05; ** p < .01; ***p < .001

Appendix C: Meta-Analysis Methodology

Biostat's Comprehensive Meta-Analysis software was used for the meta-analysis. In extracting effect size data, SRI analysts entered numeric data that met these two criteria:

- Sufficient data elements were provided to permit calculation of an effect size.
- The sample size of both the treatment group and the comparison group was 24 or larger.

According to these criteria, seven projects provided estimates of effect sizes, or quantitative data for both treatment and comparison groups so that an effect size could be calculated. Most organizations included data disaggregated by campus, course or implementation of the same courses. This finer-grained information was used in the meta-analysis when it was available. Each campus, course, or implementation was treated as a subgroup of the study in the meta-analysis.

The type and amount of data provided in project reports varied by grantee. Data provided in these formats were used in the meta-analysis:

- ***Sample size and event rate for treatment and control group:*** the most common format for course completion. The N and course completion rates were used to calculate an odds ratio, which was then converted to Hedges's g.
- ***Mean standard error and sample size for treatment and control group:*** the most common format when grantees reported measures on a continuous scale.
- ***Difference in means, standard deviation and sample sizes:*** used when grantees did not provide means for treatment and control group, but provided the difference in means, standard deviation, and sample sizes, permitting calculation of Hedge's g.
- ***Cohen's d, standard error and sample size:*** some grantees provided calculated effect sizes in the form of Cohen's d. They were converted to Hedges's g for consistency.

Two meta-analyses were conducted on student outcomes: one on the course completion, a binary variable, and the other on continuous measures of student learning, such as course grade or examination score. Although odds ratios or log odds ratios are the usual ways to represent effect sizes for binary variables, all the effect sizes were converted to Hedges's g for consistency in reporting and convenience in comparing impacts on different kinds of outcomes.

Raw Data vs. Data Adjusted Based on Modeling

If the project used any kind of modeling or statistical control for potential confounding of other variables with treatment condition, SRI analysts extracted the adjusted means and standard deviations based on the modeling to calculate effect sizes. Otherwise, the raw data were used in the meta-analysis.

Meta-Analysis Models and Unit of Analysis

Because the interventions varied substantially from one project to another, a random-effects model was applied in the meta-analysis.

For the overall meta-analysis, the unit of analysis was project rather than individual course implementation (specific course, campus, and/or semester of implementation). The software calculated the effect size for every course implementation, collapsed the implementations within each project, and then aggregated the effects at the project level.

The overall meta-analyses for binary and continuous variables were followed by tests of the courseware features listed in Exhibit 12 as possible moderator variables to ascertain whether effect sizes varied by each of these courseware features. Since some projects included multiple courses with different features, the course implementation was treated as the unit of analysis in the moderator variable analyses. This choice also increased the power of the tests for moderator variables since it increased the sample size from 6 projects to 62 (course implementations).

Appendix D: Additional Project Information

Grantee	For More Information:
Pathways Project, Carnegie Foundation for Advancement in Teaching	Clyburn, G. M. (2013). Improving on the American Dream: Mathematics pathways to student success. Change: *The Magazine of Higher Learning, 45*:5, 15-23. Available at http://dx.doi.org/10.1080/00091383.2013.824346 Silva, E., & White, T. (2013). *Pathways to improvement: Using psychological strategies to help college students master developmental math.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Strother, S., Van Campen, J., & Grunow, A. (2013, March). *Community College Pathways: 2011-2012 descriptive report.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.Available at http://www.carnegiefoundation.org/sites/default/files/CCP_Descriptive_Report_Year_1.pdf Van Campen, J. V., Sowers, N., & Strother, S. (2013, December). *Community College Pathways: 2012-2013 descriptive report.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Available at www.carnegiefoundation.org/.../pathways/CCP_Descriptive_Report_Year_2.pdf
NROC/DevMath	Martin, T., & Brasiel, S. (2014, February). *STEM Action Center Technology Pilot Assessment.* Logan, UT: STEM Action Center, Utah State University. Available at http://stem.utah.gov/about-stem/pilot-project/ The NROC Project. (no date). N*ROC Developmental Math Case Studies.* Available from http://nrocmath.org/products/higher-ed/developmental-mathematics/#case-studies_tab The NROC Project. (2013). *NROC Developmental Math Fact Sheet.* Available from http://nrocmath.org/cms/wp-content/uploads/2013/11/NROC-Developmental-Math.pdf
NCAT/Changing The Equation	National Center for Academic Transformation. (no date). *Redesigning Mathematics: Increasing Student Success at a Reduced Cost.* Available at http://www.thencat.org/RedMathematics.htm National Center for Academic Transformation. (no date). *How to Redesign a College Course Using NCAT's Methodology.* Available at http://www.thencat.org/Guides/AllDisciplines/TOC.html National Center for Academic Transformation. (no date). NCAT Redesign Scholars Program. Available at http://www.thencat.org/RedesignAlliance/ScholarsProgram.htm Twigg, C. A. (no date). *Improving learning and reducing costs: Program outcomes from Changing the Equation.* Available at http://ncat.org/Mathematics/CTE/CTE_Lessons.html Twigg, C. A. (2011, May-June). *The math emporium: Higher education's silver bullet. Change, the Magazine of Higher Learning.*
OpenStax College/ Rice University	*OpenStax*. May 2014 Newsletter. Available at: https://openstaxcollege.org/news
University of Maryland MOOCs, ITHAKA S+R	Griffiths, R., Chingos, C., Mulhern, C., & Spies, R. (2014, July). *Interactive Online Learning on Campus: Testing MOOCs and Other Platforms in Hybrid Formats in the University System of Maryland. ITHAKA S+R.* Available at http://sr.ithaka.org/research-publications/Interactive-Online-Learning-on-Campus Griffiths, R. (2013, October). *MOOCs in the Classroom? ITHAKA S+R.* Available at http://www.sr.ithaka.org/blog-individual/moocs-classroom ITHAKA S+R. (2013, October). *Interim Report: A Collaborative Effort to Test MOOCs and Other Online Learning Platforms on Campuses of the University System of Maryland.* Available at http://www.sr.ithaka.org/research-publications/testing-moocs-and-other-online-learning-platforms-campuses-university-system
University of Toronto MOOC	Open UToronto. (no date). *MOOC Research and Evaluation.* Available at http://www.ocw.utoronto.ca/mooc-research-and-evaluation/ Rolheiser, C. (2014). *Hatch, match, and dispatch: Examining the relationship between student intent, expectations, behaviours and outcomes in six Coursera MOOCs at the University of Toronto.* Available at http://www.moocresearch.com/wp-content/uploads/2014/06/MOOC-Research-InitiativeRolheiser9167-May-Report-.pdf

SRI Education

333 Ravenswood Avenue
Menlo Park, CA 94025
Phone: 650.859.2000

www.sri.com/education

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

BILL & MELINDA GATES foundation