Lessons from
Five Years of Funding
DIGITAL COURSEWARE

Executive Summary

September 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.

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

Cover photo used with permission, courtesy of Montgomery College.

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

Executive Summary

Barbara Means

Vanessa Peters

Ying Zheng

September 2014

As the Bill & Melinda Gates Foundation's Postsecondary Success strategy enters its fifth year of learning technology investments, it is a good time to take stock of what has been learned and to draw implications for future investments. The foundation asked SRI Education to review the major courseware-related projects in the Postsecondary Success portfolio and provide an independent synthesis of findings.

The foundation identified the 12 major postsecondary courseware-related projects in Exhibit 1 as sufficiently completed to contribute to SRI's review. Three of the projects were actually sets of multiple grants or subgrants addressing a common goal. In total, the courseware investments reviewed by SRI involved 137 courses and represented approximately 90% of the foundation's financial investment in postsecondary courseware over the last five years.

Primary data sources for this synthesis and review were final and interim reports submitted by the 12 projects and their subgrantees, interviews with principal investigators, and project-related research articles and additional data provided in response to SRI's request. Proposals, RFPs (requests for proposals), and project websites provided additional background information.

SRI analyzed the features of the 137 different courses developed or evaluated through these projects and performed a quantitative meta-analysis of student outcomes for those projects that provided the data needed to estimate the impact of the project's courseware.

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 (massive open online courses) as they are known 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.

Portfolio Description

Features of the courseware developed, implemented, and (in some cases) evaluated by these projects were coded using a set of online and blended learning features adapted from the conceptual framework provided by Means, Bakia, and Murphy.[1] That framework organizes features of online and blended learning into three dimensions: context of use, instructional and technology design, and implementation practices.

---

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

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 content 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.
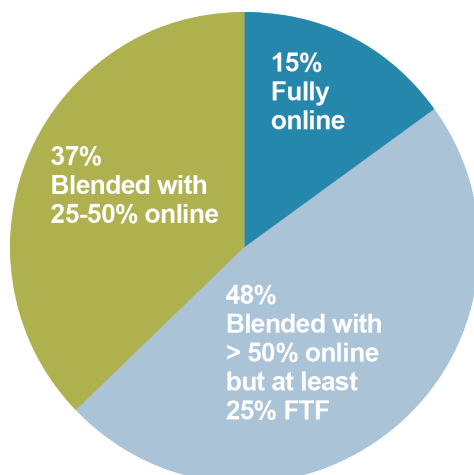
The foundation's emphasis on improving college success for underrepresented minorities, low-income students, and first-generation college goers was reflected in the context for which their grantees designed and implemented instructional software.

• 60% of courseware in the Postsecondary Success portfolio was used primarily in community colleges, and

• 78% explicitly targeted low-achieving students (notably, students requiring remediation before enrolling in a college-level math course).The courseware's role within the college course varied: Approximately 44% of the courseware efforts were redesigns of existing courses, 12% were whole new courses, and 12% involved developing digital resources to supplement existing course materials. The remaining courseware efforts were coded as supports for the process of redesigning a course in a blended format (11%), learning analytics or early warning systems embedded in existing courseware or a learning platform (4%), or online supports for peer learning (4%).

In terms of subject matter, 57% of the courses dealt with mathematics, 16% with one of the sciences, 12% with English or humanities, and 15% with one or more other subjects. These findings are all consistent with the Postsecondary Success team's identification of community college and developmental (remedial) mathematics courses as key areas requiring improvement in order for the foundation to meet its goal of doubling the college completion rate for low-income students.

In terms of implementation practices, just 15% of the courseware in the Postsecondary Success portfolio was intended for fully online implementation without a co-located instructor. The rest was designed with the expectation that it would be used in a blended or hybrid model, in which students did some of their learning online and some with a faculty member or teaching assistant. The expectation for the proportion of time spent online is shown in Exhibit 2.

The human instructors in these courses were not expected to engage in extensive synchronous interaction with students online or to foster learning through online interactive activities in which the learning content is emergent as students interact with each other and the instructor.[2] In the great majority of implementations

Exhibit 2. Proportion of Time Spent Online in Postsecondary Courseware Projects

15%
Fully online

37%
Blended with 25-50% online

48%
Blended with > 50% online but at least 25% FTF

Note: FTF = face to face

[2] See Stephen Downes, Learning Networks: Theory and Practice, *International Conference on Methods and Technologies for Learning,* Palermo, Italy, 2005; Margaret Riel and Linda Polin, Online Communities: Common Ground and Critical Differences in Designing Technical Environments (pp. 16-50) and Thomas M. Schwen and Noriko Hara, Community of Practice: A Metaphor for Online Design (pp. 154-178), in Sasha A. Barab, Rob Kling, and James H. Gray (Eds.), *Designing for Virtual Communities in the Service of Learning.* Cambridge, UK: Cambridge University Press, 2004.

(89%), the classroom instructor did not interact with students in any significant way online. Students' dominant activity when working with the courseware was either solving problems and answering questions (68% of projects) or listening and reading (22%).

Description of the Evidence Base

Three of the Postsecondary Success courseware projects in this review were technical assistance efforts that did not plan to provide evidence that a courseware-based intervention improved student outcomes (Exhibit 1). In addition, several of the projects that did plan to measure student impacts as a long-term goal were not yet at that stage of work when this review was conducted (spring 2014). For these reasons, only 7 of the 12 projects provided student outcome data that SRI could use in its review of findings. However, three of these seven were programs with subgrants to other organizations, and a number of projects included multiple implementations of one or more courses, so that in total some kind of student outcome data were available for 105 courseware implementations.

For these 105 implementations,

- Student course completion rate was by far the most commonly reported outcome, being documented for 90% of cases.
- Performance on a post-assessment or course final was reported by 44%.
- Final course grade was the outcome reported by 17%.
- Accumulated college credits were reported by one project.

In addition to these student academic and learning outcomes, many projects reported the number of students or campuses using their courseware and instructor and student satisfaction data. Two projects, Changing the Equation and ALMAP, and some of the NGLC subgrants (e.g., Kaleidoscope and BioBook projects) reported institutional or student cost savings.

Quantitative Synthesis of Impacts on Course Completion and Student Learning

Meta-analysis is a technique for synthesizing the results from a series of studies quantitatively. It has the advantages of

• being 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;

• producing synthesized effect estimates with considerably more statistical power than individual studies; and

• allowing an examination of differential effects related to different courseware features (moderators) such as a hybrid or fully online course design.

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 differ 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.

Using project and subgrantee reports to the foundation, journal and web-published articles describing results of courseware implementations, and outcome data provided in response to requests for more data, SRI extracted information permitting the calculation of an effect size for 105 courseware implementations. 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),[3] analysts performed two separate meta-analyses: one using course completion rates and the other using outcome measures such as a grade or post-assessment score.
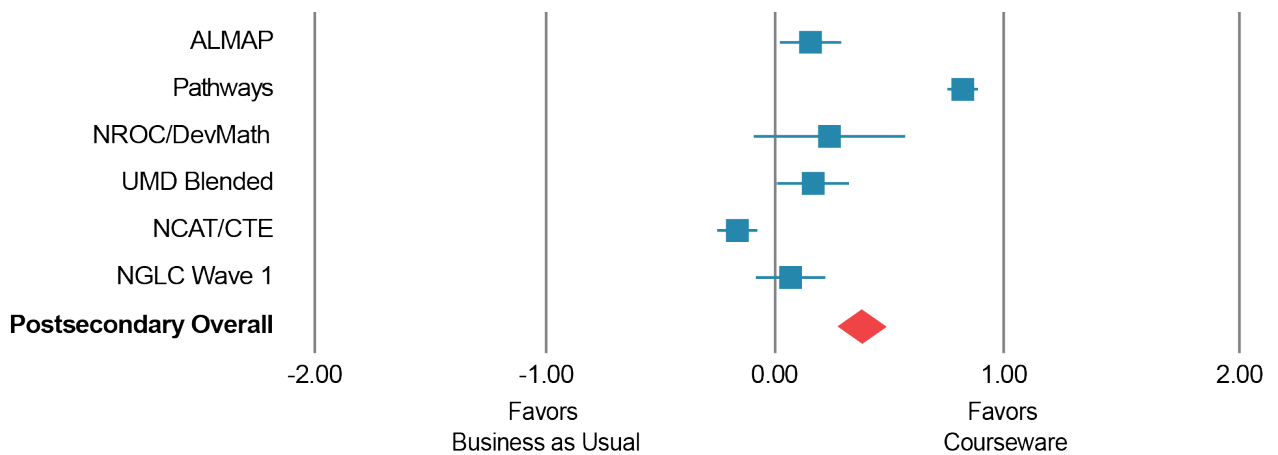
Course completion data from 94 course implementations were used in the first meta-analysis. When examined at the level of individual course implementations, there were 55 cases of no impact on course completion rate, 22 cases of a significant negative impact, and 17 cases of a significant positive impact. Analysts then aggregated across individual course implementations to produce an average effect size for each of the six projects with course

---

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

completion data as well as an overall average for the Postsecondary Success portfolio. As shown in Exhibit 3, the Postsecondary Success projects as a whole had a moderately positive effect on course completion rates. The overall average effect size of .37 standard deviation units is equivalent to almost doubling the likelihood that a student who starts a course will finish it. However, this average effect estimate was influenced strongly by a single project with a large positive impact (Pathways). If the average Postsecondary Success courseware impact on course completion is estimated with the Pathways data excluded, the average effect estimate is very close to 0. As seen in Exhibit 3, five of the six projects with effect estimates for course completion had effect sizes that were close to 0. Certainly the observed impacts on course completion have been much more modest than the dramatic improvement envisioned in early articulations of the Postsecondary Success strategy.

Exhibit 3. Courseware Impacts on Course Completion Rate

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.

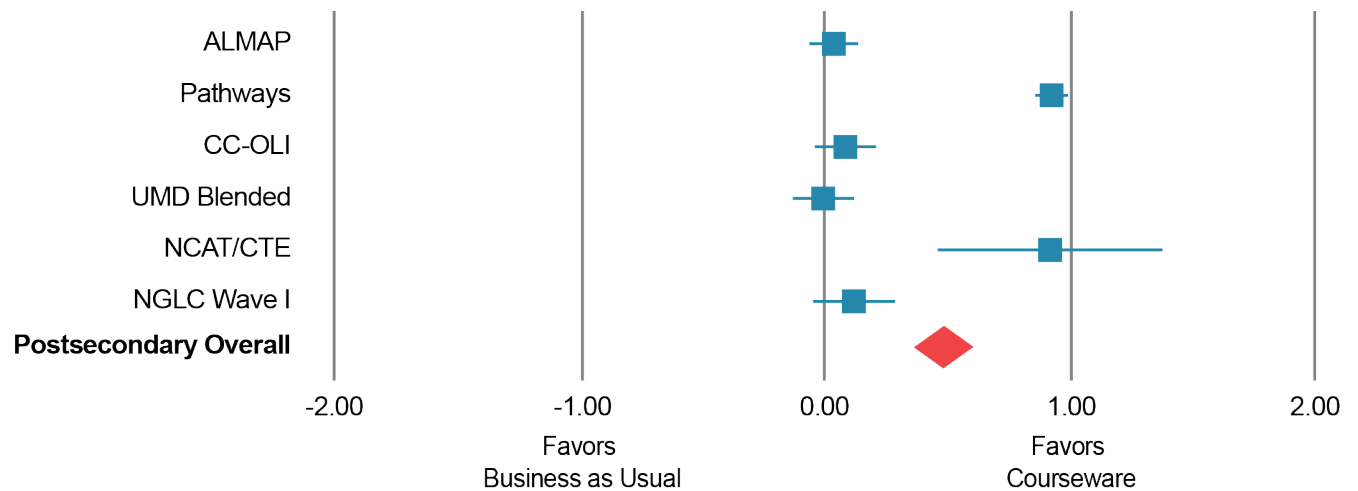## Exhibit 4. Courseware Impacts on Continuous Variables

Exhibit 4 illustrates the meta-analysis results for learning outcomes (such as final examination score or course grade) available for 62 course implementations. For these learning outcomes, **_the overall Postsecondary Success impact was moderately positive—an effect estimate of .47._** 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 large positive impacts, with effect estimates of .89 and .92, respectively. The four other projects with learning outcome data had effect estimates indistinguishable from 0.[4]

Both of the projects with large positive impact estimates for learning outcomes involved a complete redesign of developmental math courses, but they used very different instructional designs. Colleges working with NCAT in the Changing the Equation (CTE) project adopted a mastery learning approach (in which a student does not advance to new content until high proficiency on current content has been demonstrated) and used math software such as ALEKS or MyMathLab for a significant part of course instructional time. These changes had the dual goals of improving student outcomes and reducing costs. Relative to students in the prior versions of the colleges' developmental math courses, students in 29 of 34 NCAT/CTE-redesigned course implementations who finished the course had significantly better learning outcomes than students completing prior versions of the math course. However, the percentage of enrolling students who finished the course in a single semester dropped in many of these cases (as reflected in NCAT/CTE's average negative impact on course completion rates shown in Exhibit 3).

_____

[4] Although the NGLC project did not have a significant impact on average, its effect size was based on estimates from 13 different subgrants, several of which had significantly positive impacts of moderate size (California State University, Northridge and Missouri Community College Consortium).

The two Pathways courses, Statway and Quantway, used software to provide students with problem-solving practice but did not use mastery learning. Both of these courses involved a redesign of the developmental math sequence, making it possible for a student not only to get through developmental mathematics, but also to earn a math credit toward graduation within a single year. Pathways courses also placed a strong emphasis on addressing socioemotional issues related to math learning and academic persistence, and the courses had positive effects on course completion rates as well as on learning outcomes.[5]

Analysis of Features Moderating Courseware Impacts

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 greater interest than the average effect size per se. SRI used the 62 learning outcome effect estimates that went into the meta-analysis of learning outcomes to explore potential moderating variables. For each coded courseware feature, analysts divided the courseware implementation codes into logical groups with the goal of having a minimum of 10 effect estimates for the smallest subgroup. In those cases where this criterion could be met, SRI tested the feature as a moderator variable.

The courseware implementation sample had enough feature variability to permit testing the influence of four aspects of the context in which the courseware was implemented: field of use (4-year versus 2-year college), students' academic preparation level (low versus medium or high), subject matter (mathematics versus other), and courseware role (whole course or course redesign versus other). All four features were associated with differences in student learning outcomes:

• 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.

• Learning effects were greater in community colleges than in 4-year colleges.

• Impacts were greater for mathematics than for other subject areas and significantly positive for math but not for science courses (other subject areas had too few effect estimates to warrant a statistical comparison).

---

[5] Elena Silva and Taylor White, *Pathways to Improvement: Using Psychological Strategies to Help College Students Master Developmental Math,* Carnegie Foundation for the Advancement of Teaching, 2013.

Six aspects of instructional and technology design (predominant student role, pacing, dominant online pedagogy, individualized learning path, modality, and technology system/platform type) were tested as possible moderators of courseware impact. Three of these six features were significantly associated with the magnitude of the courseware's learning impact:

• Courses in which students' dominant role when working online was solving problems or answering questions had more positive effects than those where the dominant online activity was listening or reading.

• Course implementations using individualized pacing had more positive impacts than those with class-based or a mixed form of pacing.

• Adaptive learning technologies demonstrated larger learning effects than nonadaptive ones.

There was enough variation in documented implementation strategies to test two features as potential moderator variables, and both were found to be significantly associated with the magnitude of the impact on learning:

• 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.

• Courses with 100 to 299 students had larger learning effect estimates than courses with either larger or smaller class sizes.

The features our analysis identified as significant moderators of courseware learning impacts are summarized in Exhibit 5.

Several limitations of the available data set and of meta-analytic techniques should be noted. Tests for the significance of moderator variables identify features associated with differences in the dependent variable (in this case, student learning measures), but they do not demonstrate that the moderator variable caused the outcome. Both the moderator variable and the outcome measure could be products of some other factor or combination of factors. Further complexity is introduced by the fact that moderator variables may be strongly correlated with each other. In the Postsecondary Success courseware data set, many projects involved the redesign of math courses to incorporate online practice, individualized instruction, self-pacing, and mastery learning. Each of these features was associated with greater impacts, and there were not enough cases to disentangle their respective influences because they so often occurred together.

Exhibit 5. 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.

A second limitation of the meta-analyses reported here is that most of the outcome data were provided by the courseware developers and institutions rather than by an independent, objective evaluator. Few projects were evaluated using the most rigorous experimental design: There were very few cases of experiments in which students were assigned at random to use the courseware or take the usual version of the course in a way that ensured that the two groups of students whose outcomes were being compared were equivalent at the outset. Some of the projects measured the characteristics of the students entering the redesigned course and the conventional version of the course (such as prior GPA) and provided learning outcome estimates that had been corrected statistically for any differences, but these were only 17 of the 62 learning impact estimates. To provide insight into the extent to which the impacts

summarized here might be inflated by differences between the students experiencing the courseware and those in traditional versions of courses, SRI tested statistical control for prior differences as a moderator variable using the same approach applied to courseware and implementation features. Learning outcome effect size estimates were significantly larger for those projects that did not use statistical controls.

Lessons Learned and Gaps in the Knowledge Base

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. Although there was an overall positive impact of Postsecondary Success funded courseware projects on course completion rates, only a single project (Pathways) had a positive effect of consequential magnitude. The course completion data suggest that the Postsecondary Success strategy was overly optimistic about the power of circa 2010 learning technology to improve course completion rates by itself.

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, for example, 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.[6] 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%.[7]

---

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

[7] *Ibid*, p. 171.

Postsecondary Success Portfolio Review: Executive Summary

To date, the Postsecondary Success courseware portfolio has provided relatively little evidence regarding 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.

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.[8] The latter assumption did not bear up well in the NGLC Wave I evaluation.[9] 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.[10]

---

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

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

[10] *Op. cit.,* Means, Bakia, & Murphy, 2014, pp. 165-177.

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.

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 was the measurement of the number of college credits earned within 2 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.

Recommendations for Future Courseware Investments

Drawing on the lessons learned and knowledge gaps identified here as well as our years of experience studying learning technology, SRI offers nine 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. 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 their courseware 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. 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). 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

Postsecondary Success Portfolio Review: Executive Summary

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.

Courseware funders should put more emphasis on third-party, independent evaluations of impact. 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 are pressured to complete course development and be ready to implement at the start of the designated academic term. Moreover, when projects do combine evaluation activities with course design and development, there is an understandable tendency to pick the outcome measures that make 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 invested 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. 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, 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.

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.

SRI Education

BILL&MELINDA
GATES *foundation*