RESEARCH REPORT

Evaluation of Adaptive Learning in Statistics (ALiS)

Testing an Online Adaptive Learning Platform at Nine Postsecondary Institutions in Maryland

Amanda Briggs *Theresa Anderson* *Semhar Gebrekristos* *Alphonse Simon*
Alice Mei

November 2019

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.

Contents

Acknowledgments

This report was funded by the Bill & Melinda Gates Foundation, through a grant to Ithaka S+R. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

We would like to express our gratitude to the following people who worked on the project or offered their input in the development of this report: Matt Chingos, David Blount, and Casey Simmons at the Urban Institute; Catherine Bond Hill (Ithaka S+R) and William E. "Brit" Kirwan (Transforming Post-Secondary Education – Math (TPSE-Math) who were the Principal Investigators on the Gates Foundation grant which made the project possible; Richard Spies, Lisa Krueger, Jenna Joo, and Martin Kurzweil, at Ithaka S+R; MJ Bishop and Nancy O'Neill of the William E. Kirwan Center for Academic Innovation at the University System of Maryland; Benny Johnson and Murray Kimball of Acrobatiq (now VitalSource); Annika Many of edBridge Partners, LLC; and administration and instructional staff at the University of Maryland College Park, Montgomery College, Anne Arundel Community College, Community College of Baltimore County, Harford Community College, Wor-Wic Community College, University of Maryland Baltimore County, Frostburg State University, and Towson University.

Executive Summary

The Adaptive Learning in Statistics (ALiS) Project tested the implementation of an adaptive learning platform developed by the company Acrobatiq[1] in introductory statistics courses at nine higher education institutions in Maryland.[2] Acrobatiq served as a textbook replacement inside and outside the classroom to introduce course concepts, and all ALiS sections had in-person class meetings. Instructors used class time to lead the students through activities to reinforce statistics concepts as part of a flipped classroom approach.

This report describes study results from the 2017/18 academic year. The Urban Institute evaluation compared course satisfaction and learning outcomes of students who completed introductory statistics in class sections with the ALiS approach against the outcomes for students enrolled in “business as usual” introductory statistics sections. Specifically, outcomes assessed in this study included course grade, passing with a C or better (the inverse of the commonly-used “DFW rate”), statistics competency (based on standardized assessment), and student satisfaction with the course.

Findings

The ALiS intervention was effective at four-year colleges relative to business as usual. Students at community colleges in ALiS sections did not show significant differences in any of the measures of performance relative to students in business-as-usual sections, but they were clearly less satisfied with the course than their peers in other sections. When all students are aggregated in the analysis, the results showed very modest improvements in student grades and statistics competency, no difference in passing with a C or better, and a lower level of satisfaction.

Results varied considerably across colleges, particularly between two-year and four-year institutions:

- Students at four-year colleges experienced small but statistically significant positive impacts across the board, with gains in their course grades (0.16 points on a 4.0 scale), probability of passing with a C or better (3.8 percentage-points), statistics competency (0.11 standard deviations), and satisfaction (0.26-point increase on a five-point scale).
- At two-year colleges, there were no significant effects on academic outcomes for students enrolled in ALiS sections (as measured by course grade, passing with a C or better, or statistics competency), and they were substantially less satisfied (-0.36 points on a five-point scale) than students in traditional sections.

- Across all colleges and both semesters, when pooling the institutions together and weighting colleges equally, students enrolled in ALiS sections experienced a very modest gain in grade (0.08 points on a 4.0 scale) and statistics competency (0.08 standard deviations), but were less satisfied with the course than their peers enrolled in traditional courses (-0.09 points on a five-point scale). There were no significant effects on students' likelihood of passing the course with a C or better.
- The ALiS initiative intended to help students who have worse outcomes in introductory statistics courses. Subgroups identified to test the effectiveness of the ALiS course in reducing gaps in success rates between well-prepared students and those less well prepared included first-generation college students, Pell grant eligible students, and students who had prior developmental education experience. When examining aggregate impacts of the ALiS intervention on these three student subgroups (e.g., first-gen vs. not-first-gen students) across all colleges across both semesters, there are no significant differences in impacts on course grades or likelihood of passing with a C or better. The ALiS intervention did not significantly affect gaps in grades or pass rates, but it seems to have increased the gap in statistics competency between groups. It also left to Pell Grant-eligible students and students who had prior developmental education experience feeling less satisfied with the course.

Descriptive analyses of instructor and student survey results provide additional context for these findings. Surveys asked how ALiS instructors structured in-person class time, including the amount of time spent on certain activities, course pedagogy, instructor dashboard use, and student perceptions of the course. Survey analyses shed light on some of the reasons the intervention may have proved most promising for students in four-year colleges—for example, we found that instructors at the four-year institutions had more time to engage with the Acrobatiq adaptive learning platform prior to the start of the semester. However, confounding factors exist. Data provided by Acrobatiq revealed the degree to which the tool was implemented with fidelity, such as the amount of time pilot students spent in the online platform and the percent of activities completed. Two-year students on average spent more time in the platform than their four-year peers while completing roughly the same percent of activities, and two-year instructors were more likely to use real-time student data provided by the Acrobatiq platform to better understand students' progression through the course.

Overall, these findings are consistent with the growing body of evidence on technology-enhanced instruction, which has demonstrated that, while students tend to fare similarly in their academic outcomes, they are generally less satisfied with the courses than their peers. Our findings generally support this trend, with two-year college students, students with prior development education experience, and those who are Pell-Grant eligible particularly dissatisfied with the ALiS technology intervention. The improved academic and satisfaction outcomes across the board for students enrolled in ALiS sections at four-year colleges are promising and should be explored further. A better understanding of what makes an intervention successful at one type of institution versus another could help program designers develop interventions that are better suited to students in each implementation context.

As the use of adaptive learning tools in and outside the classroom continues to grow, additional research is needed to identify which aspects of online learning tools, pedagogy, and course structure tend to have the biggest impact on student outcomes and how those different elements interact in ways that are effective for both instructors and students. This includes a better understanding of what kinds of support can be provided to promote a positive experience for students who complete courses that have an adaptive learning component and what kinds of training and support are needed for instructors to effectively use the data and other resources available through the adaptive learning platform. Additional research is particularly needed at community colleges. By identifying the most important factors for student success, instructors and institutions looking to implement technology-enhanced instruction at scale might be able to move beyond the “do no harm” principle and affect positive change for students.

Introduction

The Adaptive Learning in Statistics (ALiS) project tested the implementation of an adaptive learning platform developed by the company Acrobatiq (a developer of online learning programs) in introductory statistics courses at participating higher education institutions in Maryland, beginning as a pilot in fall 2016.[3] The study scaled from two participating institutions in fall 2016 to nine institutions in spring 2018. This report describes study results from the fall 2017–spring 2018 academic year.

The project was supported with funding from the Bill & Melinda Gates Foundation. Ithaka S+R, a not-for-profit organization dedicated to improving higher education through increased use of digital technologies, directed the project. The ALiS project team[4] facilitated college participation and designed professional development support for instructors implementing the new course design. The project was supported by the University System of Maryland's Kirwan Center for Academic Innovation, which played a convening role across project partners. The Urban Institute has been the project's independent evaluation partner since 2016.

The ALiS intervention consisted of a technology component to deliver course content and a pedagogical approach designed to reinforce concepts learned outside of class during in-person class meetings. The ALiS intervention could be considered a "blended classroom approach", where online activity is mixed with classroom meetings.[5] Introductory statistics course content was delivered via the Acrobatiq online learning platform[6], which served as an interactive textbook replacement. The Acrobatiq platform was considered "adaptive" in various ways. Students completed formative and summative learning assessments in the Acrobatiq tool, and Acrobatiq used data from these assessments to generate a "learning estimate" for each student within each course learning objective. That learning estimate informed the content of quizzes at the end of each module, designed to assess students' level of understanding. Acrobatiq data also fed into dashboards that instructors could use to monitor student progress and adapt their teaching to address challenges students faced in the course.

A unique element of the ALiS intervention that varies from most "blended learning" models is that face-to-face time with instructors was not reduced for most students in ALiS sections.[7] As a result, instructors were strongly encouraged to develop course pedagogy that complemented students' efforts in Acrobatiq outside of class. The primary pedagogical model was a flipped classroom, in which students first encountered content through assigned work outside of class and completed hands-on concept reinforcement activities during class meeting times. Details of the intervention appear later in this report.

The Urban Institute evaluation compared satisfaction and learning outcomes of students who completed introductory statistics in class sections with the Acrobatiq tool against those of students enrolled in "business as usual" introductory statistics sections. Because the ALiS project team advised participating instructors to implement Acrobatiq using a flipped classroom approach, the evaluation also describes the pedagogical approach used in the classroom by ALiS instructors.

We begin by describing the important role that technology-enhanced instruction is playing in the adaptation of traditional educational models toward more scalable and sustainable options for students. We then describe our evaluation methods and research process, summarize our findings, and discuss implications for future research.

Evidence on Technology-Enhanced College Instruction

Previous evaluation studies sponsored by Ithaka S + R have found null or small, positive impacts of technology-mediated interventions on student learning. Findings of research efforts to date support the "do no harm" principle. That is, student course outcomes were not adversely affected by participation in the course, but they were not significantly improved either. These studies include:

- a series of randomized experiments testing hybrid statistics course on six public university campuses (Bowen et al., 2012),
- a nonexperimental study of redesigned courses using elements of Massive Open Online Courses on public university campuses (Chingos et al., 2016), and
- a series of randomized experiments testing if adaptive online learning can be used to enhance math preparedness (Griffiths, Chingos, and Mulhern., 2015).

Other studies have focused more specifically on the impact of adaptive learning technology on student learning. A study of a 2015-2016 pilot at the University of California involving the use of adaptive technology found that at all three campuses where the intervention was implemented as intended, student course performance outcomes were positive, including for some at-risk populations (University of California 2016). Another quasi-experimental study by SRI International on the early implementation outcomes of nine adaptive courseware products in courses across 14 colleges found that the impacts of adaptive courseware on student outcomes, cost, and student and instructor satisfaction varied by institution type, course subject and modality of instruction. The authors note that future studies of the effectiveness of adaptive courseware should consider the implementation context and how different approaches to instruction can affect learning outcomes (Yarnall et al. 2016).

While evidence is still emerging in this area, further investigation is needed to understand the importance of course structure and pedagogical approach on student outcomes. For example, fully online programs make postsecondary education more accessible to students, but these programs can also create academic barriers to student success. A recent review of the literature cites evidence that student outcomes, as measured by course completion rates and satisfaction, are worse in fully online courses and degree programs. However, compared to fully online courses, hybrid/blended courses[8] do not have the same negative outcomes. Rather, they have been found to "generally yield similar or improved outcomes relative to standard classrooms" (Protopsaltis and Baum 2019, 32).

Hybrid/blended learning coursework—which combines an in-person instructional experience with online learning components—has been implemented and piloted at numerous two- and four-year colleges. These courses employ a unique learning modality that seeks to combine the flexibility provided by an online class with in-person interactions with an instructor (Potter 2015). By design, blended learning courses are designed to incorporate more student-instructor interaction than fully online coursework. Peer-reviewed studies confirm the significance of student-instructor interaction as a key component of quality that leads to higher student satisfaction and achievement. A meta-analysis of 74 studies in 2009 concluded that stronger interaction and greater engagement between students and instructors is associated with improved achievement and stronger outcomes (Bernard et al., 2009).

Perhaps because of the importance of student-instructor interaction in the classroom, technology-mediated interventions have historically been integrated into course curricula to support traditional teaching methods, involving lecture as the core method of instruction (Nielsen and Yahya 2013). However, to increase engagement with the learning process, some instructors have shifted toward a blended learning approach that involves a flipped or inverted classroom. In a "flipped classroom," students first gain exposure to course material outside of the classroom and use class time to complete activities or problem-solving exercises that allow them to reinforce the concepts they learned outside of class (Brame 2013). There is no consensus on a universal definition of this pedagogical approach, and the impact on student outcomes from this approach is not yet well understood (Chen et al., 2019 and Limniou et al, 2018). A meta-analysis on student learning outcomes in hybrid, blended, and flipped courses found that among "mixed-method" courses—those that utilize both face-to-face and online instructional methods—the only type of course that consistently improved student learning outcomes delivered content via technology and provided feedback via the instructor (Margulieux et. al. 2015). The authors of the meta-analysis termed this approach the "flipped blend".

The potential benefits of and barriers to successful implementation of technology-mediated interventions on student outcomes are ripe for further exploration. Our study adds new research to the field regarding student impacts resulting from an adaptive learning intervention to examine who benefits and why. We also incorporate a descriptive analysis to highlight how a flipped, blended classroom approach in pilot sections of the course affected students' course experience and discuss outstanding questions for future research.

The ALiS Intervention

The ALiS project team designed an intervention informed by the American Statistical Association Guidelines for Assessment and Instruction in Statistics Education (GAISE)[9] and the Dana Center Math Pathways (DCMP) model.[10] The ALiS intervention was designed to align with certain principles from the DMCP framework: to integrate strategies that support students as learners within the course and to incorporate evidence-based curriculum and pedagogy that reinforced students' positive self-perception as math learners and improved problem-solving and communication skills.[11]

ALiS also adopted a set of common learning objectives for introductory statistics. These learning objectives were previously established as part of an ongoing mathematics reform initiative in the state of Maryland led by the University System of Maryland (USM), called the First in the World Maryland Mathematics Reform Initiative (MMRI).[12] The MMRI is a collaborative effort between the public four-year USM institutions and two-year community colleges designed to help students who require developmental education more efficiently and effectively accelerate into credit-bearing postsecondary courses through the implementation of rigorous high-quality math pathways.

The theory of change underlying ALiS was that the combination of an online learning tool and complementary pedagogy during in-person class meetings would lead to meaningful differences in student outcomes in introductory statistics coursework (i.e., student satisfaction, course grade, and statistics competency). In particular, the project hoped to boost the success rates of low-income (Pell-eligible), low-skill (with developmental math experience), and first-generation college students.

The logic model in Figure 1 in describes the inputs, activities, outputs, and anticipated outcomes associated with the theory of change, and describes how the Acrobatiq tool was designed to be implemented in the classroom in the 2017–18 school year at participating institutions.

FIGURE 1

Adaptive Learning in Statistics Logic Model

Inputs

- Introductory statistics curriculum
- Technology infrastructure (wi-fi; tablets and/or computers)
- Professional development resources for participating instructors (Virtual Learning Community)
- Adaptive learning platform (Acrobatiq)
- Instructors willing to implement the tool and able to effectively engage with technology
- Adaptive online learning platform with embedded statistics curriculum provided to students at no cost
- Stipend from Ithaka to offset implementation costs and effort ($20,000 for the academic year)

Activities

Functional

- Colleges designate a lead instructor to interface with Ithaka project team and Acrobatiq
- Lead instructors and Ithaka project team work together to develop and refine statistics curriculum and recommended pedagogical approach
- Lead instructor or head of math department recruits instructors to participate in pilot
- Colleges procure access to Acrobatiq software
- Instructors receive training from Acrobatiq on how to use the platform
- Instructors receive information from Ithaka regarding the resources available for reference in the Virtual Learning Community
- Colleges identify spaces with appropriate infrastructure for blended learning/in-class activities
- The university system supports the cross-college learning network
- Ithaka, the university system, and institutions hold regular check-in calls and meetings to share best practices and address instructor concerns

Programmatic

- Students complete course curriculum outside of class using Acrobatiq software
- Acrobatiq software offers adaptive course pathways for students based on initial level of proficiency
- Teachers use student activity data in the platform and personalized learning assessments to understand if students are falling behind in coursework
- Teachers use Virtual Learning Community materials for in-class activities and other active learning resources
- Teachers implement a flipped classroom approach where they facilitate hands-on activities in class to supplement student learning in Acrobatiq

Outputs

- Students complete at least 90 percent of recommended activities and assessments using Acrobatiq
- Teachers use student activity data in the platform to intervene when students fall behind
- Students achieve mastery on learning outcomes
- Acrobatiq metadata and survey responses confirm that flipped classroom approach was implemented

Outcomes

Short-Term:

- Students perform better, as measured by course grade, course pass rate, passing the course with a C or better, and statistics competency
- Students are more satisfied with their course experience
- Instructors are satisfied with their experience teaching the course
- The gap between education-or-resource disadvantaged students and non-disadvantaged students is narrower or closes

Intermediate:

- Students are more likely to enroll in additional advanced math classes
- Colleges achieve greater consistency and efficiency in math pathways course that improves overall math achievement

Long-Term:

- Colleges find financially sustainable ways to continue using adaptive learning
- Transfer and articulation from two-year to four-year colleges improves due to curricular alignment
- Achievement gap between at-risk and traditional students is reduced
- Students have better math and statistics competency

The core ALiS intervention was for instructors to use the Acrobatiq course in lieu of a textbook as the primary curriculum for the semester. The ALiS project team and partners provided various resources to instructors to facilitate the implementation of the platform into their course structure. Those resources included the necessary technology infrastructure (e.g., computer-enabled classrooms for in-class activities), professional development resources for instructors, and a stipend from the ALiS project team to participating institutions to offset the implementation costs and effort. The Acrobatiq platform was also provided to students free of charge during the study period.

Prior to the fall 2017 semester, project partners agreed on common learning objectives and began to explore different approaches to implement Acrobatiq. The ALiS project team strongly suggested the flipped classroom approach as the preferred pedagogy for the intervention, though this was not strictly enforced. A flipped classroom in this intervention meant that instructors were encouraged to support student learning by facilitating hands-on activities in class to reinforce key concepts as students progressed through the Acrobatiq content outside of class. The ALiS project team conveyed the expectation that instructors cover 90 percent of the Acrobatiq curriculum, with students completing all related course content (reading, activities, and assessments) in the platform.

To enable colleges to share best practices across the university system, the ALiS project team designated a lead instructor to serve as a liaison between the project team and Acrobatiq, convene meetings of instructors at the college, and flag issues with course content and curriculum. Through the ALiS learning community, facilitated by the ALiS project team, lead instructors at participating institutions shared sample activities and discussed best practices. In fall 2017, at the beginning of the evaluation period covered in this report, the Kirwan Center developed a broader Virtual Learning Community using an online tool called Basecamp. The same semester, Acrobatiq added new adaptive features to the platform. The design remained consistent throughout the 2017–18 school year. During this time, Acrobatiq continued to refine the details of the course, and the flipped classroom approach became more strongly emphasized.

Instructors received training and support from Acrobatiq staff on how to integrate the technology into the classroom. The Acrobatiq platform provided instructors with detailed feedback on students' progress and success through an interactive instructor dashboard, which instructors could use to further customize the in-class student experience. Implementing the ALiS model with fidelity involved using student activity data from the platform to intervene when students fell behind on coursework, enabling them to achieve mastery across learning outcomes.

To help assess short-term outcomes, students in ALiS pilot courses were tested on the same content at the end of the semester as students in traditional (i.e., business-as-usual) course sections. If their experience in the course was successful, students—including those at risk due to developmental education experience or other factors—would perform better than their peers in traditional statistics courses at their college. Expected short-term impacts on student outcomes included: improved course grades, a higher probability of passing the course, of passing with a C or better, and improved statistics competency. Students and instructors would also be more satisfied. Taken together, these short-term outcomes would translate into improved intermediate outcomes, with students participating in the ALiS section more likely to enroll in additional advanced math classes and demonstrate improved math and statistics competency.

Beyond the impact on individual students, the hope was that colleges participating in the ALiS project could identify financially sustainable ways to continue using adaptive learning approaches in introductory statistics and math education coursework beyond the end of the grant. Furthermore, if a standardized curriculum across two- and four-year colleges could be agreed upon, project partners surmised that Acrobatiq might improve transfer and articulation outcomes for students from two-year to four-year institutions.

The outcomes described in Figure 1 represent the best case scenario, in which the tool is implemented with fidelity and the instructors flip the classroom per the guidelines provided by the ALiS project team each semester. Guidelines for a flipped classroom were as follows: *"Class time is mostly dedicated to working individually or in groups to apply content that was presented online outside of class using adaptive software. Most large group discussion or lectures are delivered via the computer; however, some class time is dedicated to clarifying topics or extending a lesson."* As with any new education intervention, particularly a project designed to be implemented across different institutional contexts, all elements of the model were not implemented with fidelity across all institutions. Later sections of the report describe these deviations in more detail.

Project Participation

Seven colleges participated in the ALiS demonstration in the fall 2017 semester, and nine colleges participated in spring 2018. The colleges represented a mix of two- and four-year public institutions. Colleges were encouraged to enroll as many students as possible in the study. The institutions and student enrollment in each semester is summarized in Table 1 and Table 2. Overall, 1,493 students participated in the study during the fall 2017 semester (621 at two-year colleges and 872 at four-year colleges) and 2,315 students in the spring 2018 semester (1,049 at two-year colleges and 1,266 at four-year colleges).

TABLE 1

ALiS Pilot Enrollment, 2017–18 School Year

Two-year public institutions

	Fall 2017		Spring 2018	
College	**Pilot**	**Traditional**	**Pilot**	**Traditional**
Anne Arundel Community College	52	68	48	51
Community College of Baltimore County	101	78	201	200
Harford Community College	74	58	86	111
Montgomery College	89	101	49	50
Wor-Wic Community College	0	0	121	132

Source: Institution and math department records

TABLE 2

ALiS Pilot Enrollment, 2017–18 School Year

Four-year public institutions

	Fall 2017		Spring 2018	
College	**Pilot**	**Traditional**	**Pilot**	**Traditional**
Frostburg State University	93	91	95	63
Towson University	0	0	224	118
University of Maryland–Baltimore County	98	93	119	121
University of Maryland–College Park	271	226	294	232

Source: Institution and math department records

Instructor and Student Characteristics

Instructors

As shown in Table 3, 45 instructors participated in the ALiS evaluation over both semesters. Two-thirds of these were at two-year colleges and one-third were at four-year colleges. There were slightly more traditional than pilot instructors overall. Because instructors could teach both a traditional and pilot section, the totals in the table are less than the sums of the rows.

TABLE 3

Instructor Participation, by Institution Type

2017–18 school year

	Two-year	Four-year	Total
Pilot	22	9	**31**
Traditional	27	11	**38**
Total	30	15	**45**

Source: List of participating instructors and corresponding registrar data provided by project lead at institution each semester
Note: Because the study design for the ALiS project involved matched pairs of the same instructor teaching both pilot and traditional sections of the course, the number teaching pilot and traditional sections by institution type is not mutually exclusive.

There were 20 returning ALiS instructors across both semesters (4 in fall 2017 and 16 in spring 2018). We designated instructors as "returning" if an instructor taught a pilot section with the Acrobatiq tool in a previous semester.

Overall, nearly half (45 percent) of all ALiS instructors across both semesters had prior experience teaching with the Acrobatiq tool. All pilot instructors were experienced. Table 4 shows the median years of teaching experience in mathematics or statistics at the college level and the median years of teaching introductory statistics at their college.

TABLE 4

Pilot Instructors' Median Years of Experience, by Institution Type

Teaching experience	Two-year	Four-year
College-level mathematics/statistics at any institution	14	9.5
Introductory mathematics/statistics at current institution	8	5.5

Source: List of participating instructors and corresponding registrar data provided by project lead at institution each semester

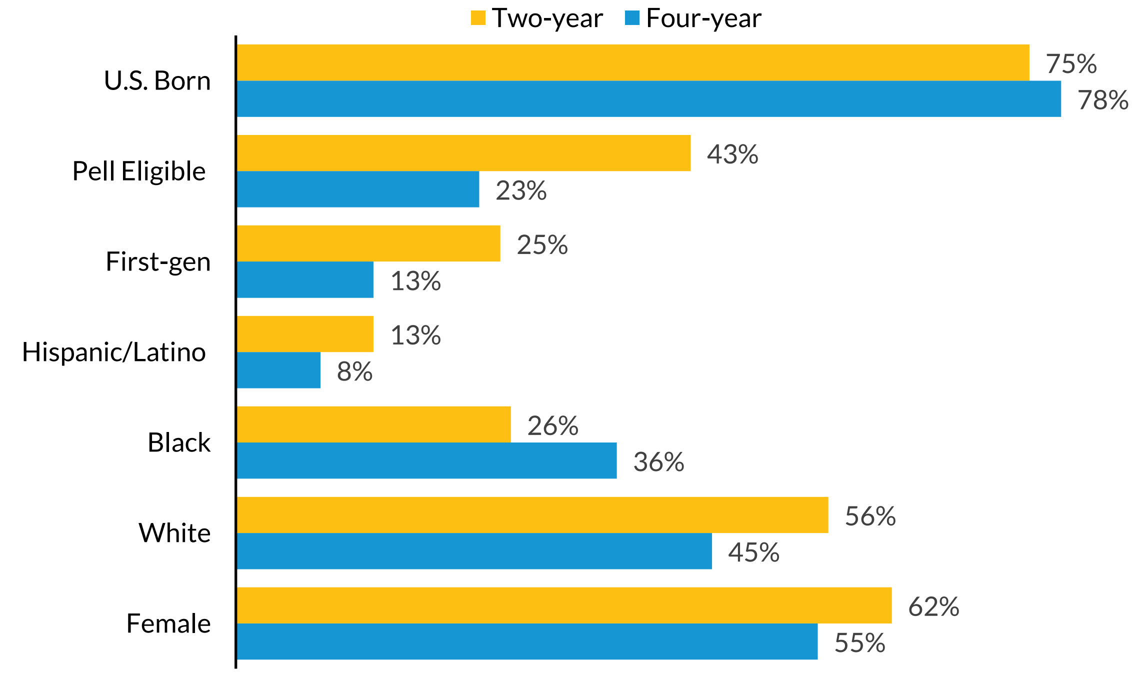
Students

Over the course of the 2017–18 school year, 3,808 students participated in the study. Student characteristics and experiences reported in figure 2 and figure 3 on page 19 are from baseline and institutional data, based on the most reliable source for each indicator. As shown in the figures, students at both two- and four-year institutions came from diverse backgrounds, but there were notable differences by institution type. Relative to students at four-year colleges, students in two-year colleges were more likely to be Pell eligible, first-generation, female, married with kids, employed during the semester, or attending school part time.

Students were also more likely to have had prior experience with developmental education coursework. The mean age of students at two-year colleges was 22.8 years old, while the mean age at four-year colleges was 20.5. Percentages represent the percent of non-missing responses for each category.

FIGURE 2

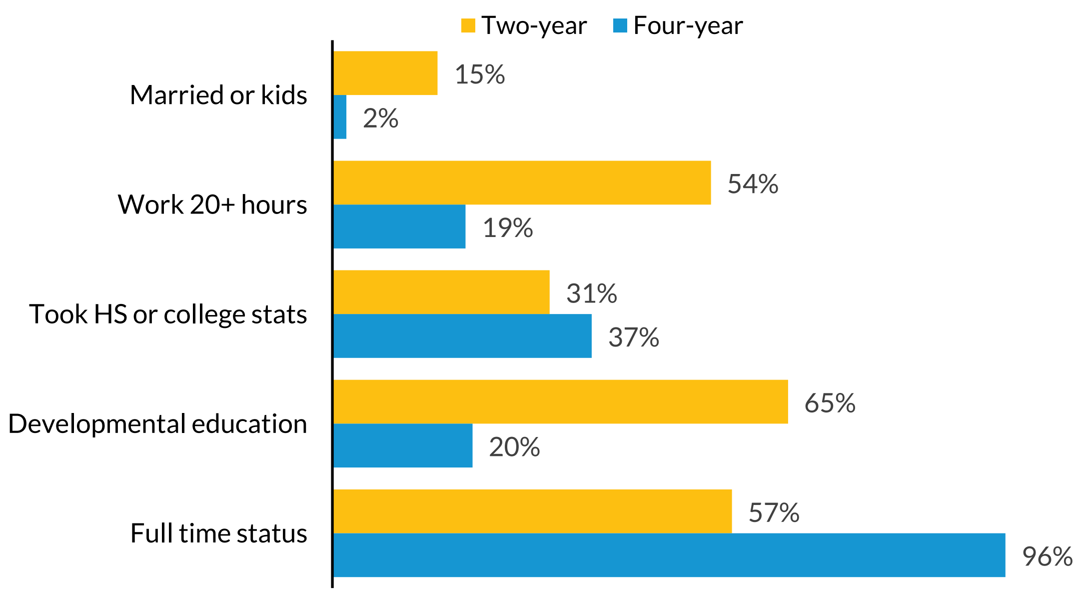
Student Characteristics, by Institution Type



Source: Institution records and student baseline survey data

FIGURE 3

Student Experience, by Institution Type



Source: Institution records and student baseline survey data

Evaluation Design

The Urban Institute led the evaluation in collaboration with the ALiS project team and the institutions to design a study that would be feasible and as rigorous as possible during the 2017–18 academic year.[13] The final approach used instructor-matched pairs to estimate the regression-adjusted impact of the ALiS intervention on pilot student outcomes compared with traditional course student outcomes.

Research Questions

The ALiS project team and the Urban research team developed a list of questions to define the scope of the study:

- Are there meaningful differences in student outcomes and completion rates for students who use the Acrobatiq tool in their introductory statistics courses versus more traditional teaching methods?
- Are there meaningful differences in student outcomes and completion rates that vary by institution type, instructor experience, and student population served?
- Does pedagogy or use of the Acrobatiq tool partially explain student academic outcomes? How so?
- What can be learned about the implementation and evaluation of online learning approaches that can be leveraged for future research?

Outcomes and Subgroups

The student outcomes of interest for the evaluation largely align with those in the Adaptive Learning in Statistics Logic Model (see):

- Course grade (4.0 scale)
- Passing course with a C or better (inverse of DFW rate, the percentage of grades of D, F, or W)
- Statistics competency (based on standardized assessment)[14]
- Satisfaction with the course (five-point scale)

In particular, this intervention intended to improve outcomes for nontraditional students who were low-income (Pell eligible), low skill (with developmental math experience), and first-generation college students–groups that otherwise struggle in introductory statistics.[15]

Sample and Data Collection Activities

SAMPLE

Table 5 summarizes the number of institutions, students, and matched pairs that participated in the study during each semester and over the entire 2017–18 school year. With the addition of two colleges in the spring semester, the overall study participation grew substantially.

TABLE 5
Study Participation

	Fall 2017	Spring 2018	Total
Colleges	7	9	
Matched pairs	20	27	47
Instructors	27	36	45
Sections			
Pilot	20	35	55
Traditional	20	29	49
Total	40	64	104
Students			
Pilot	778	1,237	2,015
Traditional	715	1,078	1,793
Total	1,493	2,315	3,808

Source: Institution records

SECTION SELECTION AND MATCHED PAIRS

In collaboration with the Urban study team, each institution designated pilot (treatment) and traditional (comparison) courses that would participate in the study each semester. Pilot sections used the Acrobatiq adaptive learning tool along with recommended pedagogical practices, to varying degrees. Traditional sections implemented the introductory statistics curriculum in its existing form. Pilot and traditional courses were selected based on willingness of the instructor to participate in the pilot and whether there were similar nonpilot courses scheduled (ideally taught by the same instructor, but also of comparable meeting length, time of day, and campus).

One pair of instructors at Montgomery College who had concurrently scheduled introductory statistics courses agreed to participate in random assignment in fall 2017. Random assignment, which is a lottery-like process that assigned half of the individuals to the treatment group (an ALiS section) and half to the control group (a non-ALiS section) proved difficult to replicate at other institutions because there were no other instances of concurrently scheduled identical courses. Furthermore, in assessing findings over the course of the 2016–17 pre-pilot year and 2017–18 pilot year, there was limited reason to believe that student characteristics across pilot and traditional course sections varied meaningfully at baseline. Rather, instructor characteristics–including prior experience teaching the course–preferred pedagogical approach, and number of years teaching, emerged as more important factors in affecting student course outcomes. This contributed to the determination that controlling for instructor effects using same-instructor matched pairs was the preferred research design.

In fall 2017, there were 20 matched pairs across seven institutions, 12 with the same instructor and 8 with different instructors. There were 27 matched pairs across nine institutions in spring 2018, 20 with the same instructor and 7 with different instructors.

DATA COLLECTION

The research team compiled information from various sources and conducted some original data collection to answer the project's research questions. Table 6 summarizes data collection activities undertaken for this project.

TABLE 6

Data Collection Activities

Instrument or source	Content	Population
Institution records	Student characteristics, instructor characteristics, student course performance	Students who enroll in traditional and pilot sections and do not opt out of data-sharing
Acrobatiq data	Student course performance, student engagement with the tool	Students who enroll in the pilot section and do not opt out of data-sharing
Student baseline survey	Student living and financial situation, statistics background, expectations for the course, and a preassessment on statistics knowledge	Students in traditional and pilot sections
Student end-of-semester survey	Activities, quality, and overall experience using computer-based tools in statistics course, and student plans for the future	Students in traditional and pilot sections
Instructor survey	Instructor teaching background, experience, and level of effort teaching an Acrobatiq pilot course	Instructors teaching pilot sections

A separate appendix containing each of the three full survey instruments can be found here. These were adapted from instruments used in Chingos et al. (2014) and customized for the project, with input from the ALiS project team and the participating institutions.[16] The baseline survey includes a small number of statistical aptitude questions pulled from the Levels of Conceptual Understanding in Statistics (LOCUS) math assessment. [17]

Quantitative Analysis Methods

To estimate impacts, we used regression analysis that controlled for matched pairs and a range of student characteristics. The outcomes and control variables are summarized in the lists below (a detailed list of covariates is available in Appendix B). In addition, each institution was equally weighted in analyses that pooled across institutions to avoid overweighting results from institutions with a larger number of participating students. Binary outcomes were estimated using linear probability modeling to ease interpretation of the coefficients. Subgroup estimates were modeled through separate regressions.

Outcome variables included:

- Grade (GPA)
- C or better (not DFW)
- Statistics knowledge (IRT ability)
- Satisfaction (five-point scale)

Control variables included:

- Demographics
- First generation
- Pell eligibility
- Developmental education history
- Prior college performance
- Standardized test scores
- Employment
- Full-time student status
- Baseline: aptitude, self-efficacy, attitudes

We also controlled for matched pair in all regressions, and all institutions were equally weighted in pooled analyses. Formally, we specified the regression as follows:

$$Y_{si} = \beta_0 + \beta_1 ALiS_{si} + \boldsymbol{\gamma}_{si} + \boldsymbol{MP}_{si} + \varepsilon_{si}$$

In this equation, *Y* is the outcome of interest for student *s* in institution ***i***. β_0 is the intercept. β_1 is the coefficient on the indicator of treatment status, ALiS. This is the coefficient of interest in the regression. $\boldsymbol{\gamma}$ is a vector of student baseline covariates. ***MP*** is a vector of matched pair indicators for each section match; this also controls for the institution, since all matched pairs are within the same institution. ε is the stochastic error term.[18] As noted, students were weighted so the total student study enrollment in each institution is equal in analyses that pool across colleges. Subgroup analyses were conducted by limiting the analyses to the relevant student subgroups, and semester-specific analyses were constrained to the semester of interest.

SURVEY ANALYSIS

Each semester, all pilot and traditional students were asked to complete two surveys, one during the first two weeks of class (baseline survey) and the second within two weeks of the end of the semester (end-of-semester survey). We asked instructors to make time in class to administer the surveys, though they could also encourage students to complete the survey outside of class if it was not possible to complete in class. As part of the baseline survey, we collected information about students' motivations for taking the course, confidence in their math skills, basic demographic information, and responses to six statistics aptitude questions. In the end-of-semester survey, students were asked to reflect on their course experience.

Pilot instructors were also asked to complete a survey describing their prior teaching experience and elements of the Acrobatiq course experience, including how they utilized the technology to monitor student progress in the course and their pedagogical approach.[19] See the Survey Instrument Appendix (linked here) for survey instruments from spring 2018.

Qualitative Research Methods

PRE-PILOT EXPLORATORY RESEARCH

This project benefited from data from a pre-pilot year from fall 2016 to spring 2017. We collected site visit data at the two institutions participating in the first year of the project. Activities included student focus groups, classroom observations, and interviews with instructors and administrators. See Appendix A for more information on research methods, interview topic guides, and findings from the pre-pilot research period.

IMPLEMENTATION TRACKING CALLS

Study activities included monthly calls between a liaison at the Urban Institute and project leads at each college. In addition, the study team had one debriefing phone call at the end of each semester with Acrobatiq to understand their team's perspective and discuss any upcoming changes or modifications to the content and resources provided to instructors in subsequent semesters.

Evaluation Findings

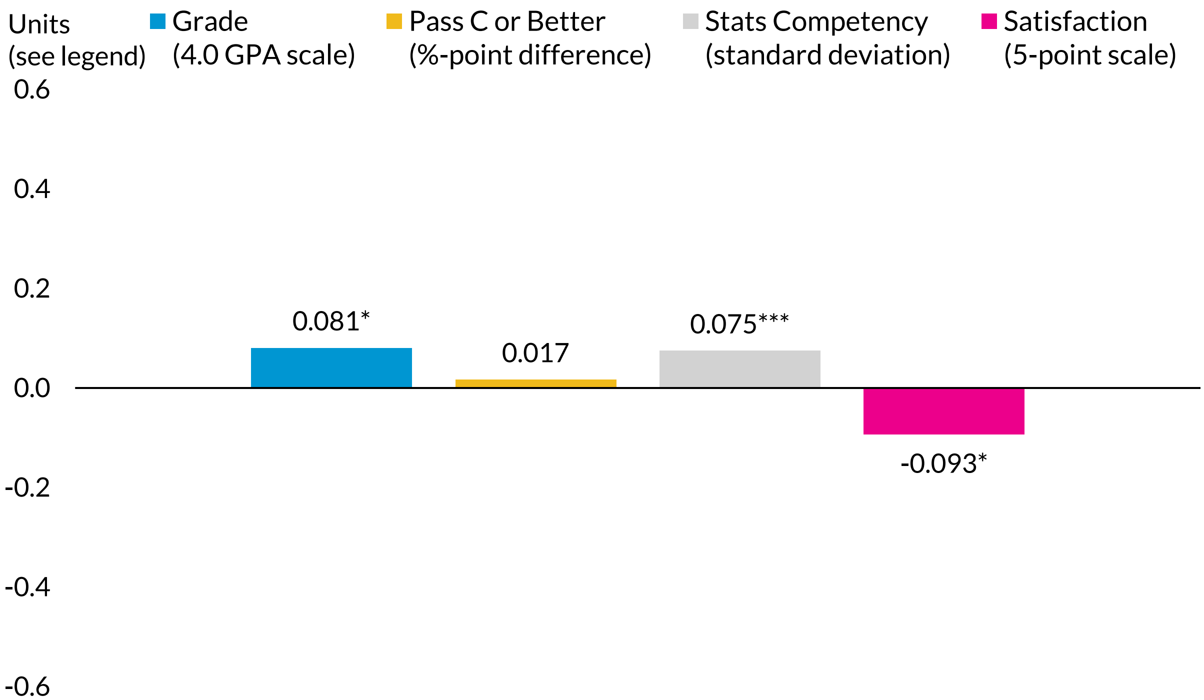
Impact Findings

The ALiS intervention improved student grades and statistics competency slightly in aggregate, but students were less satisfied. The results varied considerably across colleges, particularly between two-year and four-year institutions. Detailed coefficients appear in Appendix B.

Across all colleges and both semesters, when pooling the institutions together and weighting them equally, students experienced a modest gain in course grade (0.08 points on a 4.0 scale) and statistics competency (0.08 standard deviations), but they were less satisfied with the course (-0.09 points on a five-point scale). There were no significant effects on students' likelihood of passing the course with a C or better. Figure 4 displays these findings. The four measures are on different scales, as described in the figure notes.

FIGURE 4

ALiS impacts across all colleges, both semesters, are modest but mostly positive

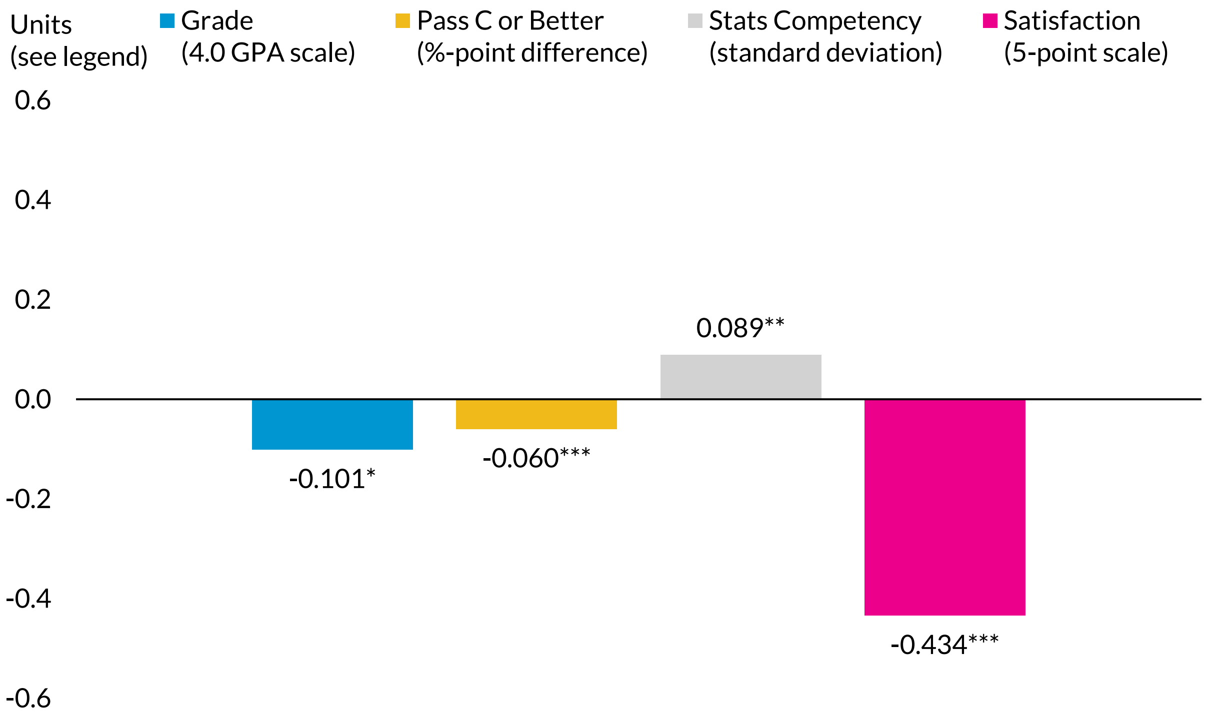


Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

One way to disentangle the results is to restrict the sample to same-instructor matched pairs. Same-instructor matched pairs control for instructor characteristics, which are important in determining student outcomes. However, two four-year institutions (University of Maryland College Park and University of Maryland Baltimore County) did not have any same-instructor matched pairs, and there were several other different-instructor matched pairs in other colleges, so these estimates do not represent impacts for all students. When restricting the analysis only to same-instructor matched pairs, the impacts are significantly negative for three of the four outcomes. Students in ALiS sections had lower course grades (-0.1 points on a 4.0 scale), were slightly less likely to pass the course with a C or better (-6 percentage points) and were less satisfied (-0.43 points on a five-point scale). However, they scored higher on the standardized measure of statistics competency (0.09 standard deviations), relative to students in non-ALiS sections taught by the same instructors. These results appear in figure 5.

FIGURE 5

ALiS impacts within same-instructor matched pairs, all colleges, both semesters, are more negative than overall impacts

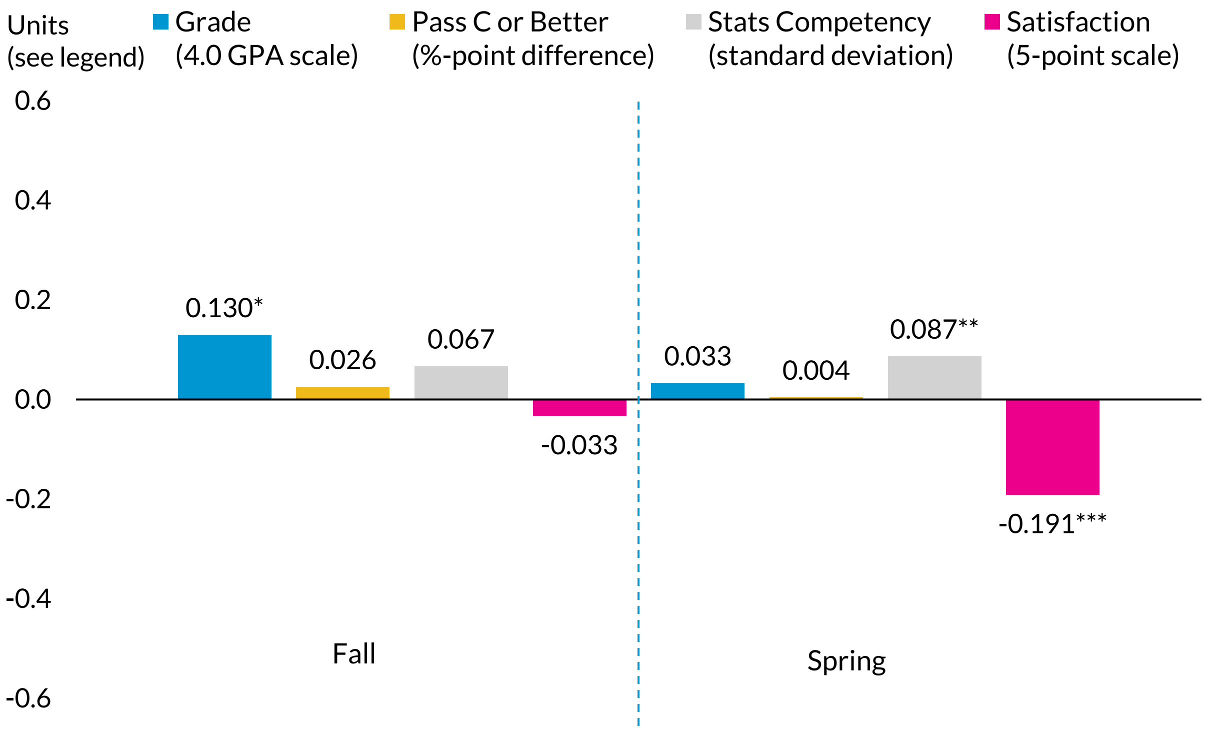


Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

There are no apparent trends in the direction or magnitude of the intervention results over the two semesters. Without limiting to same-instructor matched pairs, semester-specific impacts are quite consistent with the overall findings, as shown in figure 6. They are even more consistent when limiting the cross-semester comparisons to the seven colleges that participated in both semesters, excluding Wor-Wic Community College and Towson University, as shown in figure 7. [20]

FIGURE 6

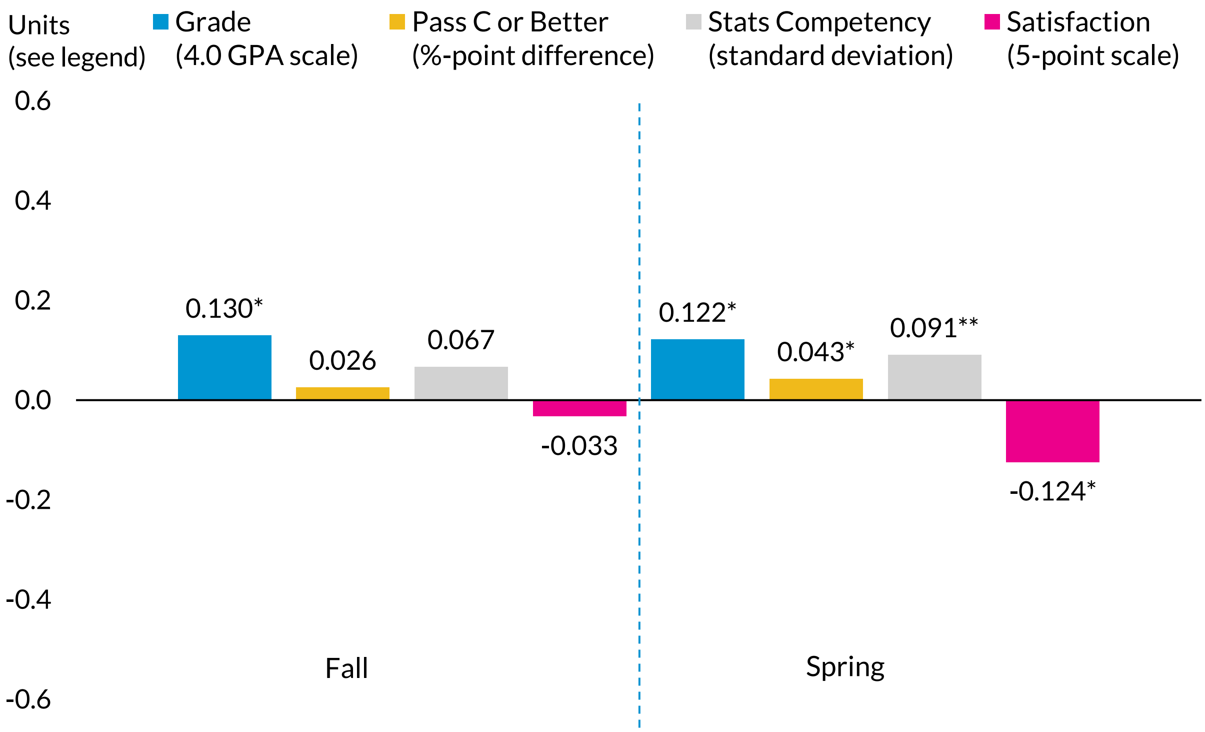
ALiS impacts across all colleges, by semester, are quite consistent



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

FIGURE 7

ALiS impacts across all colleges except Towson and Wor-Wic, by semester, are even more consistent

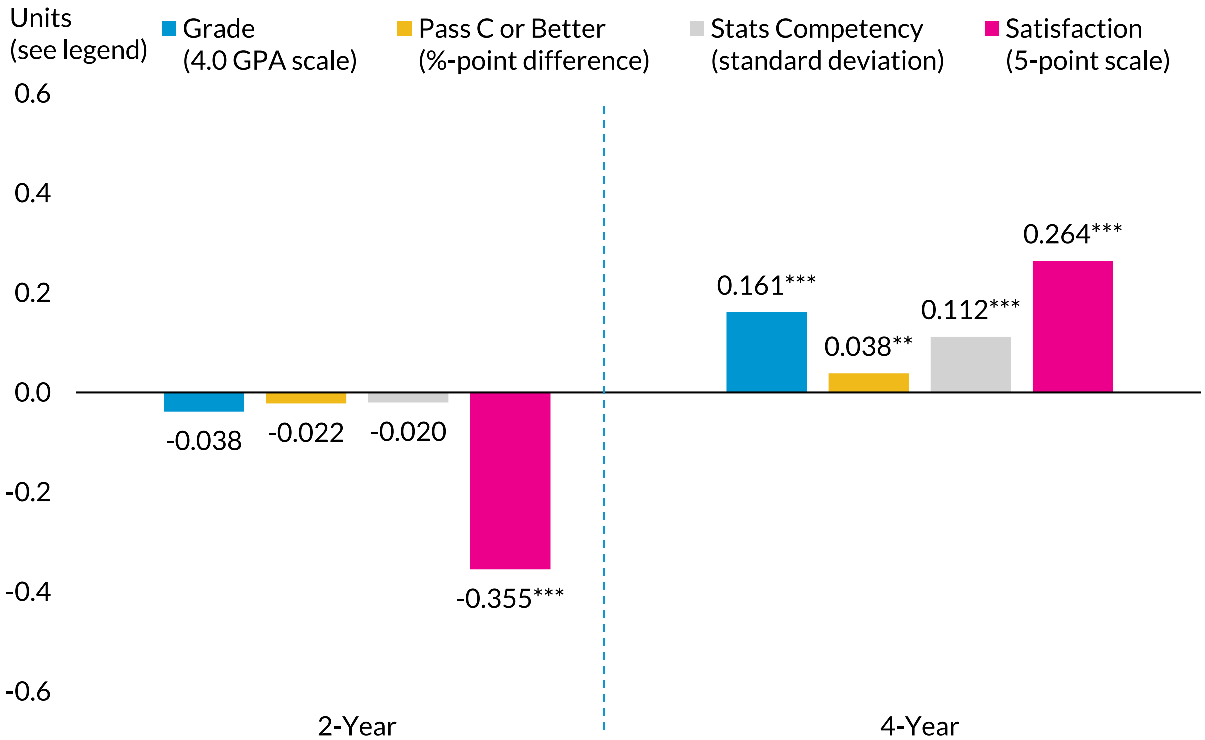


Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

We begin to see important differences when we divide the analysis by institution type. Figure 8 shows the impacts across both semesters, separated between two-year and four-year colleges. At two-year colleges, there were no significant impacts on students' academic outcomes (course grade, pass with a C or better, or statistics competency), but they were substantially less satisfied (-0.36 points on a five-point scale). Meanwhile, students at four-year colleges experienced statistically significant positive impacts across the board, with gains in their course grades (0.16 points on a 4.0 scale), probability of passing with a C or better (3.8 percentage-points), statistics competency (0.11 standard deviations), and satisfaction (0.26-point increase on a five-point scale).

FIGURE 8

Impacts across both semesters, by institution type, only reveals positive impacts at four-year colleges

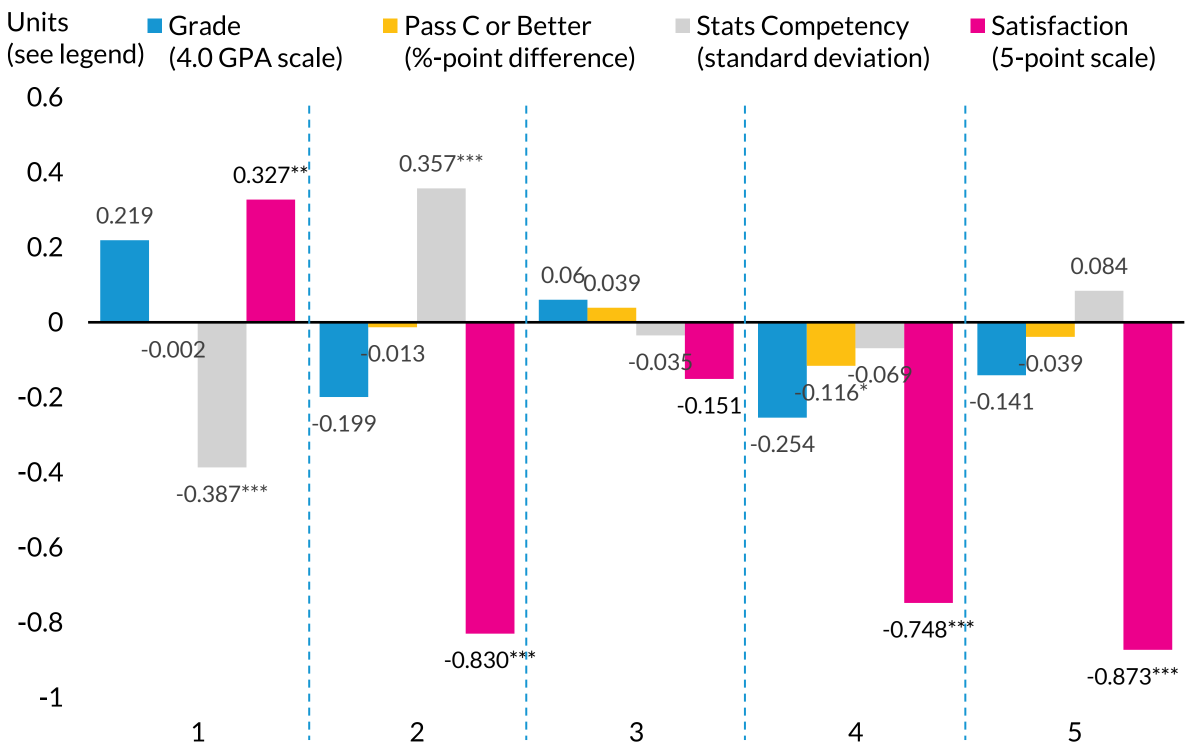


Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

While patterns of impact were generally consistent across colleges by institution type, there was some variation in outcomes at the individual institution level. Figure 9 and figure 10 display this variation (note the vertical scale differs slightly in figure 9). Each institution is de-identified but represented by a number in the graphs below.

FIGURE 9

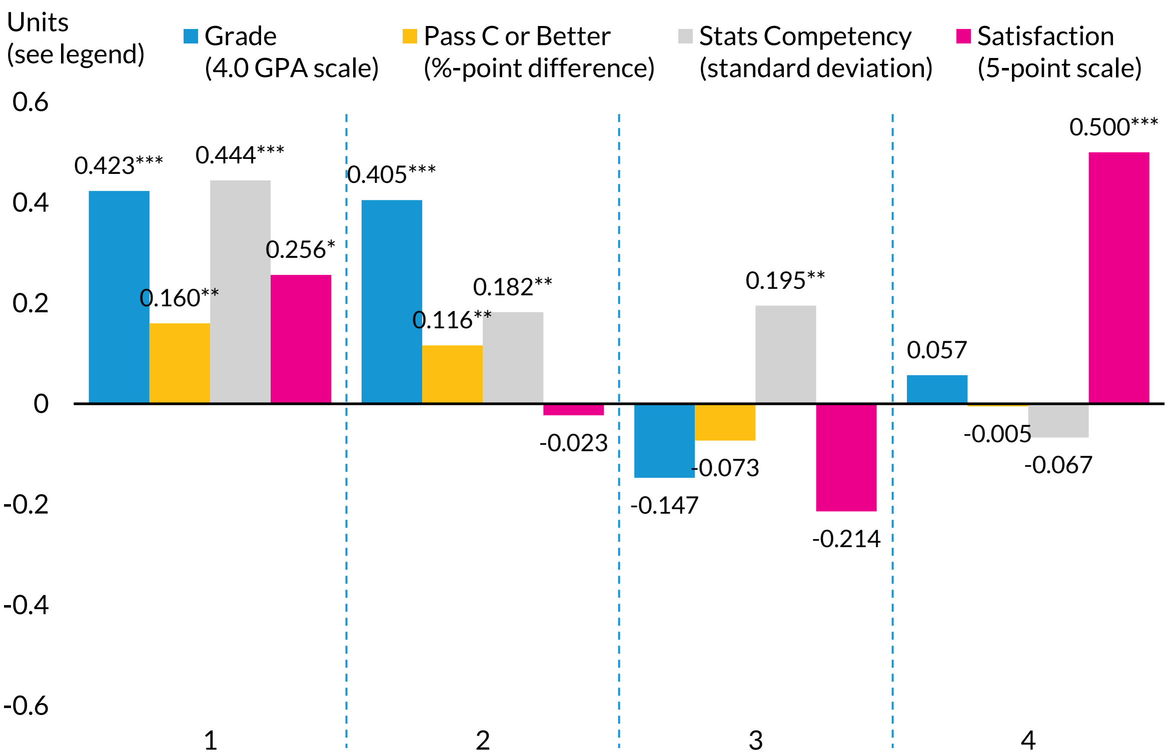
Impacts across both semesters at the individual institution level at two-year colleges are mostly neutral, with negative satisfaction



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

FIGURE 10

Impacts across both semesters at the individual institution level at four-year colleges are mostly positive, with some variation

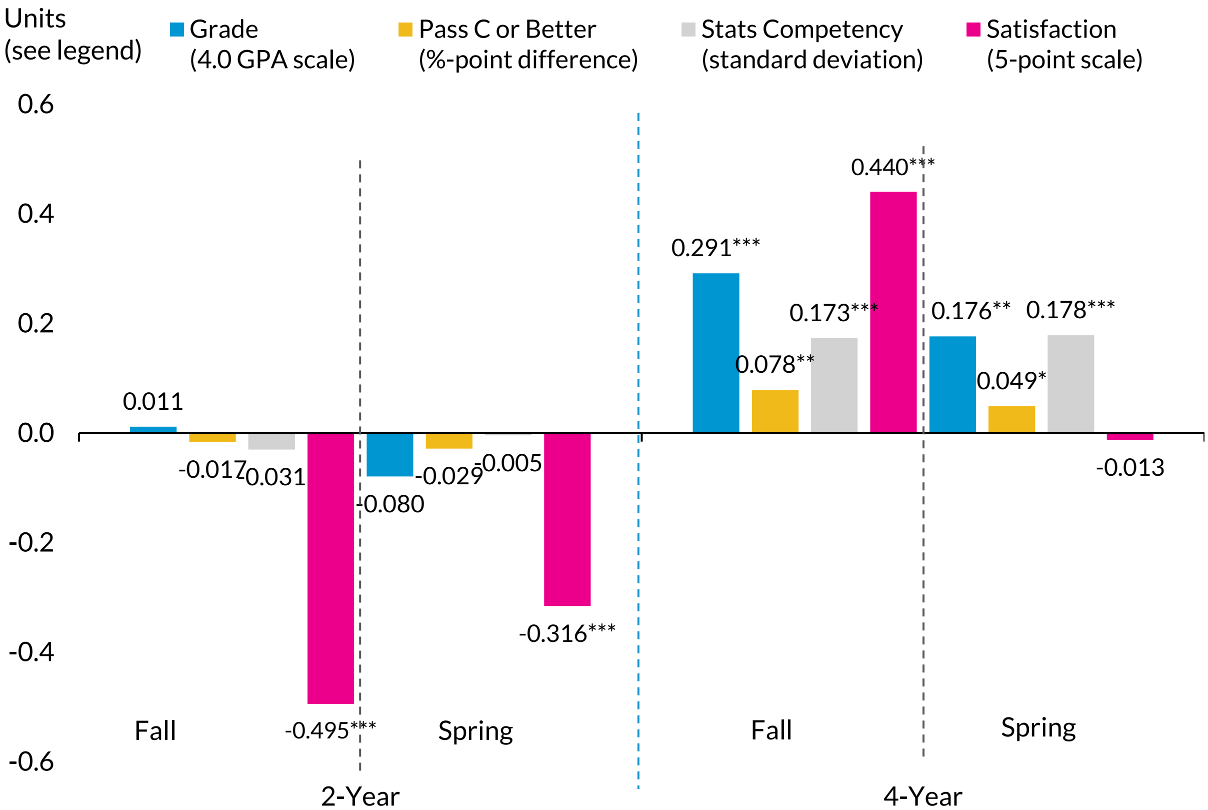


Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

Even considering the variation across colleges, patterns of impacts by institution type are consistent across semesters, as displayed in figure 11. The main variation across semesters by institution type was that the impact on satisfaction at four-year colleges was substantially larger in the fall than the spring. This may be partially explained by the addition of Towson University in the spring.

FIGURE 11

Impacts by institution type and semester show mostly consistent patterns of impacts over time



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

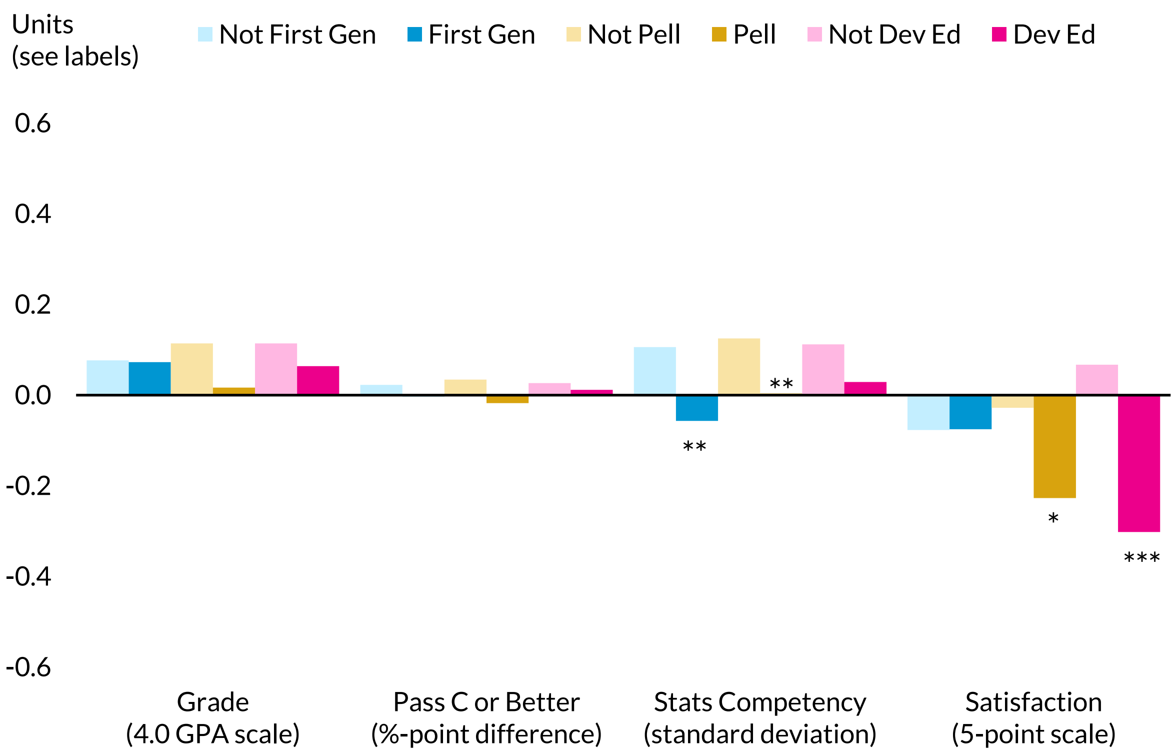
Student Subgroups

The ALiS initiative intended to help students who have historically struggled in introductory statistics—specifically, first-generation college students, Pell Grant eligible students, and students who had prior developmental education experience. Figure 12 shows the aggregate impacts on these three student subgroups across all colleges (equally weighted) and both semesters. The reader should note that the orientation of these graphs is different from the previous set of figures, with the outcomes on the horizontal axis and the bars representing student subgroups. Asterisks represent significant differences in impact estimates across subgroups (e.g., “not first gen” compared with “first gen”). These subgroup patterns look very similar across two-year and four-year colleges, so we do not disaggregate here.

Overall, there are no significant differences in treatment effects by subgroups on course grade or passing with a C or better. Students who were not first generation and students who were not Pell Grant eligible saw significantly larger gains than their counterparts in their statistics competency. Students who were Pell Grant eligible or had prior developmental education experience had significantly more negative treatment effects than those not in these respective groups. The ALiS intervention did not significantly affect gaps in grades or pass rates, but it seems to have increased the gap in statistics competency between groups and to have led to differential satisfaction.

FIGURE 12

Impacts on student subgroups across all colleges, both semesters, show that disadvantaged students generally are not harmed by participating, but they are less satisfied



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

Instructor Experience

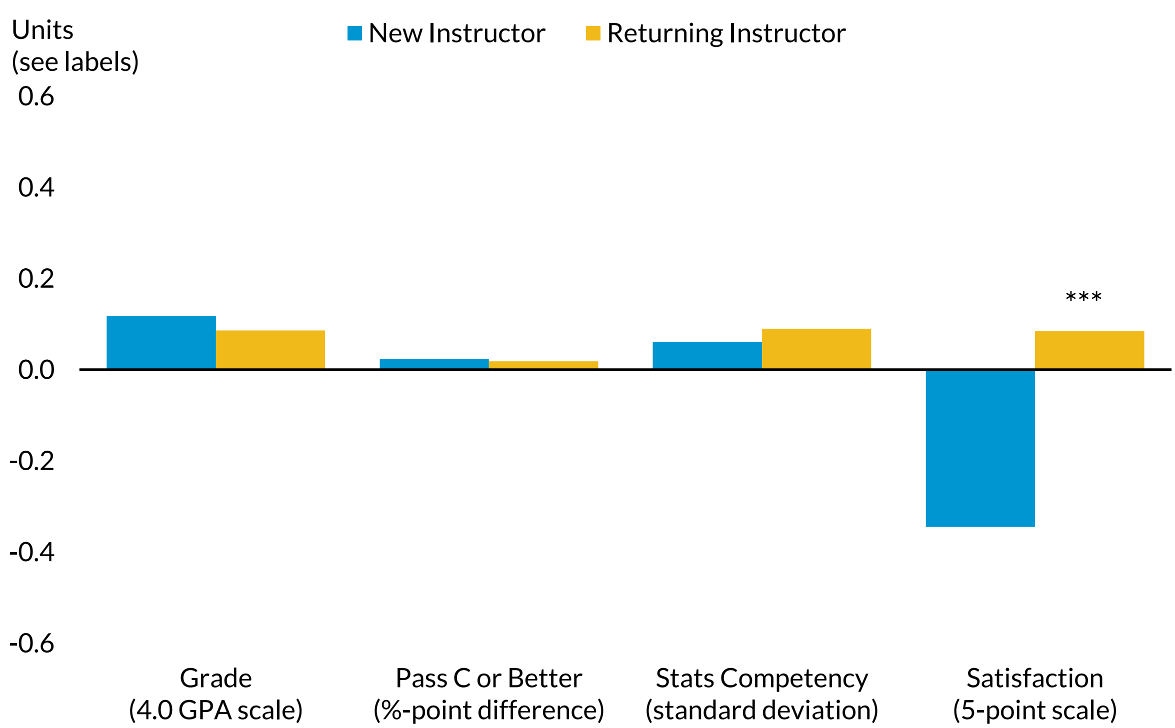
We were interested in understanding whether the ALiS intervention became more effective as instructors gained more experience using Acrobatiq and teaching with the recommended pedagogy.

Figure 13 and figure 14 show these results with two approaches. Figure 13 compares the impacts for students in ALiS sections taught by a returning instructor to the impacts for students in ALiS sections taught by a new instructor, across both semesters. Figure 14 displays the impact results in the fall and spring for students in courses that were taught by returning instructors in the spring (where the instructor taught ALiS in both semesters) to see if impacts improved with an additional semester of experience. Many of the instructors were new to ALiS in the fall and then had one semester of experience by the spring, but some at Montgomery College and University of Maryland College Park had prior ALiS experience by fall 2017, because of their participation in the ALiS pre-pilot phase.

The impacts on satisfaction for students taught by returning instructors were significantly more positive than for students taught by new instructors, as shown in figure 13. There were no significant differences in other outcomes between students of new and returning instructors. When focusing on students of the same instructors semester over semester, there does not appear to be a trajectory of improvement in impacts on student satisfaction or any other outcome, as shown in figure 14.

FIGURE 13

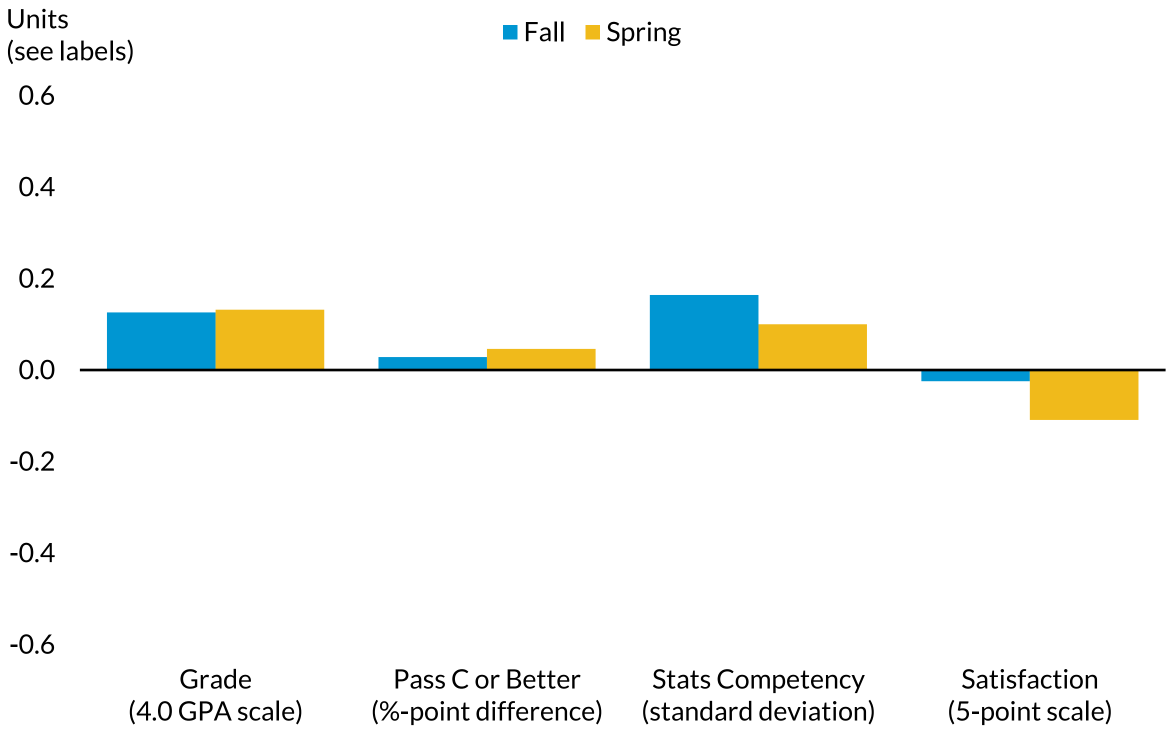
Impacts by instructor pilot experience across all colleges, both semesters, show that returning instructors have better satisfaction scores



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for Grade (out of 4.0); percentage points of probability for C or Better; standard deviation from the mean for Statistics competency; and points out of 5 for Satisfaction.

FIGURE 14

Impacts for instructors who returned in spring across all colleges, by semester, show not much change as instructors gain more experience



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. The scales for each measure are as follows: GPA points for *Grade* (out of 4.0); percentage points of probability for *C or Better*; standard deviation from the mean for *Statistics competency*; and points out of 5 for *Satisfaction*.

Elements of satisfaction

The most apparent theme in these impact results is that students in the pilot were often less satisfied than students in traditional sections, especially at two-year colleges. The end-of-semester student survey asked various questions about elements of student satisfaction, and we analyzed each element by institution type to disaggregate the overall satisfaction findings. The findings appear in figure 15. The components of satisfaction are derived from the following survey questions:

- **Satisfaction with tool:** Overall, how would you rate your experience with [tool] relative to other computer-based tools you have used in the past? (Much worse to much better)
- **Quality of instruction:** How would you rate the quality of the course meetings/lectures? (Terrible to excellent)

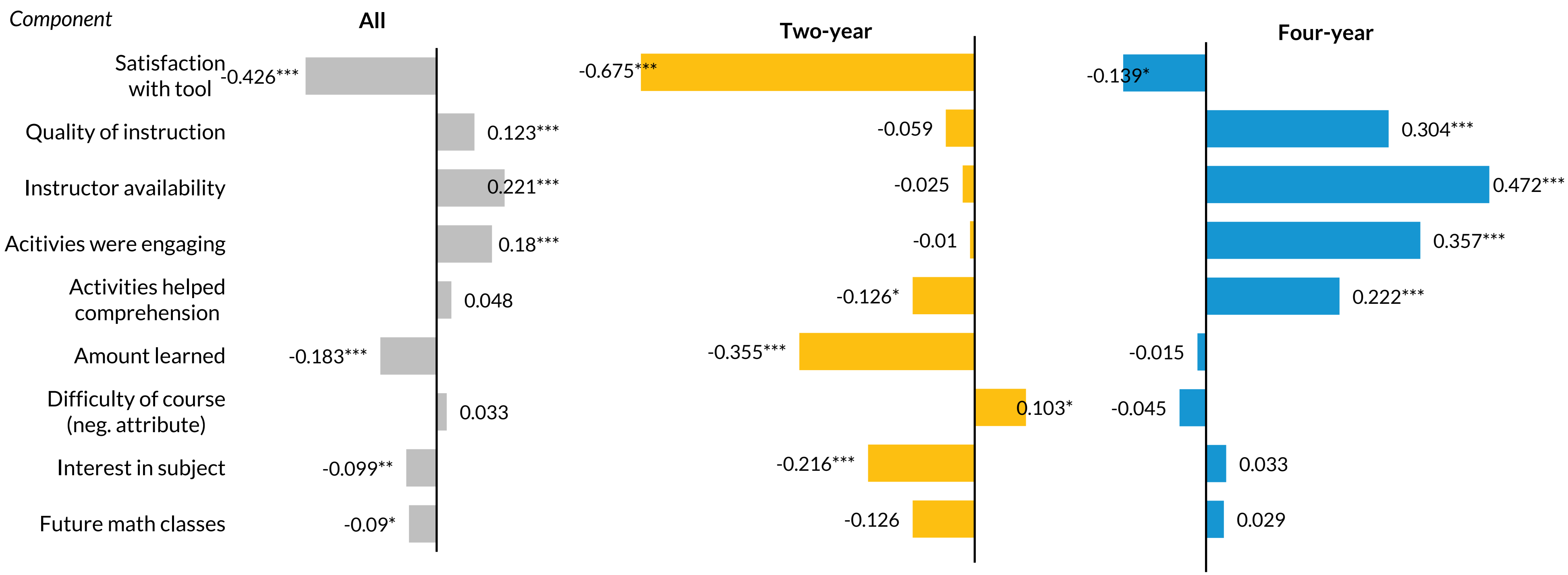
- **Instructor availability:** We want to know how this class compares with other introductory-level college classes. How available was your instructor to answer questions? (Far below average to far above average)
- **Activities were engaging:** The in-class activities were engaging. (Strongly disagree to strongly agree)
- **Activities helped comprehension:** The in-class activities helped me understand course concepts. (Strongly disagree to strongly agree)
- **Amount learned:** We want to know how this class compares with other introductory-level college classes. How much did you learn? (Far below average to far above average)
- **Difficulty of course:** How difficult did you find [course]? (Extremely easy to extremely difficult)
- **Interest in subject:** Did [course] increase or decrease your interest in statistics? (Decreased interest a lot to increased interest a lot)
- **Future math classes:** How likely are you to take more math classes? (Definitely will not to definitely will)

In figure 15, each bar represents the results of a regression on a five-point scale, with the components listed on the left-hand side as the dependent variable. Direction from the y-axis differentiates significant negative impacts (left) from significant positive impacts (right), though "difficulty of course" is a negative assessment and is therefore reverse coded. While the satisfaction findings across all students are mixed, the results diverge substantially between two-year and four-year colleges. Relative to traditional introductory statistics sections, impacts on student satisfaction with the tool were negative at both types of institutions.[21] Four-year college ALiS sections had relatively larger and more positive impacts on student satisfaction with the quality of instruction, instructor availability, the level of engagement of activities, and the amount by which activities helped comprehension.

The first three of these were nonsignificant at two-year colleges, but two-year college students thought the activities in the pilot sections were less helpful for their comprehension than traditional students. Two-year college ALiS students also felt they learned less, that the course was more difficult, and that they had less interest in the subject relative to two-year students in traditional sections. No significant differences appeared between groups for four-year college students on those outcomes. Overall, students reported they were somewhat less likely to take future math classes than their traditional counterparts, though this was not significant for either subgroup (likely because the sample size is smaller in the subgroups, so it is harder to detect significant effects).

FIGURE 15

Impact of ALiS on Components of Student Satisfaction by Institution



Note: Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01. One symbol represents 0.33-point impact on a five-point scale.

Summary of impact findings

The impact findings are small in magnitude, with more promise among students at four-year colleges than students at two-year colleges. The findings for two-year college students support the "do no harm" principle, in that students are generally not worse off in their academic outcomes. However, consistent with other research on similar interventions, students tend to report that they are less satisfied at two-year colleges. Satisfaction was positive at four-year colleges, though only in the fall semester. The following section illuminates some of the instructor and student experiences in the classroom that may help explain this pattern of impacts.

Implementation Findings

This section describes the results of the surveys of the instructors and students involved in the ALiS study. The findings provided more context for how courses were implemented across institutions and insight into the reasons for the negative student satisfaction results that emerged in the impact study.

Each semester, the Urban research team asked instructors of ALiS pilot sections to complete a survey at the end of the semester describing their prior teaching experience, how their course was structured, how they used the data dashboard in Acrobatiq to monitor student progress in the course, impressions of the Acrobatiq tool's features and functionality, and any reflections on the most promising and challenging aspects of teaching the course. The survey also asked instructors to compare their experience teaching an ALiS section with teaching a business-as-usual section of the course and their expectation for how student outcomes differed between course sections. All but one instructor (in spring 2018) completed the survey. The results of the pilot instructor survey analysis appear in this section *by pilot section* rather than *by instructor*. There were 20 ALiS pilot sections in fall 2017 and 33 pilot sections in spring 2018 included in the study.

We fielded baseline and end-of-semester surveys to students in both the traditional and pilot sections enrolled in the study. Surveying all students allowed us to collect a pretreatment and posttreatment comparison of attitudes and confidence of math and statistics, gather information about student characteristics and demographics not available in institutional data, and understand students' experiences with their online learning tool. Most students in traditional sections used an online homework tool such as MyStatLab or StatKey as a textbook replacement, allowing for comparison between students' experiences in the pilot section using the Acrobatiq tool and students' experiences in the traditional sections using other learning tools. Of the 3,808 students in the study, 3,487 completed the baseline student survey (92 percent response rate) and 2,932 completed the end-of-semester survey (77 percent response rate of all students and 87 percent of students who received a final grade in the course).

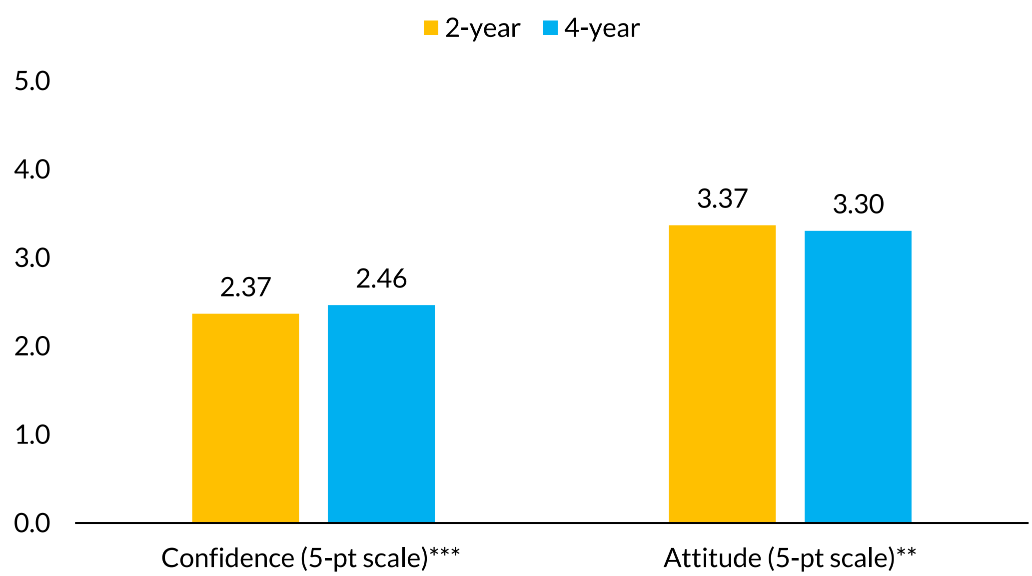
Student Attitude, Aptitude, and Confidence

Figure 16 and figure 17 show the average baseline confidence, attitude, and aptitude scores by institution type across the fall and spring semesters. Aptitude was measured using six questions from the LOCUS that were designated at the eighth-grade statistics level, while confidence and attitude were measured using a five-point original scale.[22] For each of these scores, we see significant differences between students at two- and four-year colleges. Students at two-year colleges were less confident in their understanding of concepts and ability to succeed in coursework and scored lower overall on the baseline aptitude assessment than their four-year peers. However, students at two-year colleges were more likely to agree with statements expressing positive attitudes about probability, statistics, and mathematics at the start of the semester.

FIGURE 16

Student Baseline Attitude and Confidence

Fall 2017 and spring 2018



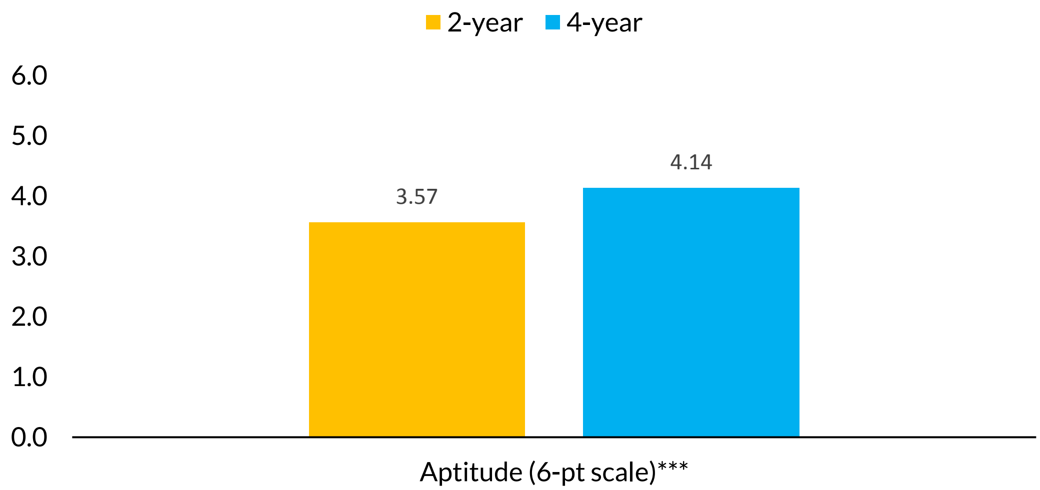
Note: These are T-tests on mean scores of confidence and attitude across institution type at baseline. Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01.

Source: Baseline Student Survey Data

FIGURE 17

Student Baseline Aptitude

Fall 2017 and spring 2018



Note: These are T-tests on mean scores of aptitude across institution type at baseline. Significance is indicated by * p<0.10, ** p<0.05, and ***p<0.01.
Source: Baseline Student Survey Data

In-Class Experience

To understand how the ALiS model was implemented, we asked about the experiences of instructors and students in ALiS classrooms, particularly the extent to which instructors implemented a flipped classroom pedagogy. We also anticipated that insight into classroom experiences might help explain the variation in impact results between two- and four-year colleges and across individual institutions.

PEDAGOGY

For a broad sense of the pedagogical approach, we asked instructors in spring 2018 about the degree to which they felt they had flipped their classroom, using the following definition:

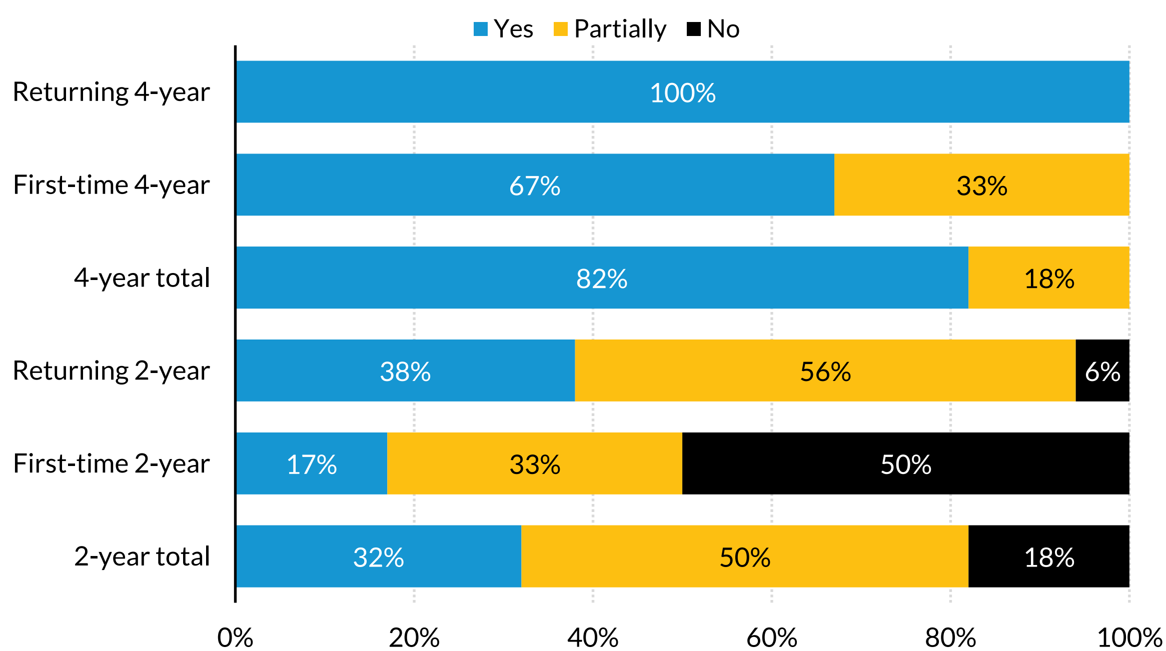
> *"The ALiS project team defines a flipped classroom as an instructional strategy and a type of blended learning that reverses the traditional learning environment by delivering the majority of instructional content online, outside of the classroom. The majority of the in-class face-to-face time is spent on targeted instruction and active learning."*

Instructors were asked to characterize their classroom pedagogy as fully flipped, partially flipped, or not flipped. As shown in figure 18, instructors reported fully flipping only 32 percent of pilot sections at two-year colleges, and 82 percent of pilot sections at four-year colleges.

Half of all first-time instructors at two-year institutions did not flip their classroom at all and nearly one in five returning two-year instructors did not flip their classrooms, while all four-year instructors indicated they partially or fully flipped their classrooms. It is possible that there is something unique about teaching in the two-year context that made community college instructors less willing to flip their classrooms, such as limited time for instructor onboarding of adjunct professors hired close to the start of the semester, instructors' response to students' stated preferences for lecture (which were documented in the pre-pilot phase), students' inability to self-teach concepts outside of class meeting time, or instructors' lack of familiarity with the pedagogy. It is difficult to discern this from the survey data, but it may be a useful avenue of inquiry for future qualitative research. Returning pilot instructors were more likely to flip the classroom than first-time pilot instructors regardless of institution type, suggesting there may be a learning curve to implementing the flipped classroom pedagogy for instructors who are new to the ALiS intervention approach.

FIGURE 18

Instructor Use of Flipped Classroom Pedagogy, Spring 2018
By returning instructor and institution type



Source: Instructor Survey Data

USE OF IN-CLASS TIME

We asked instructors and students about the use of class time and when students first encountered new statistics concepts. Differences emerged in the reported use of class time in spring 2018 across institution type and by reporting source. ALiS instructors at two-year colleges estimated they lectured 32 percent of the time, and students reported 27 percent, on average. Meanwhile, four-year ALiS instructors estimated they lectured 18 percent of the time, while students estimated closer to 26 percent. Instructors reported that they spent the remaining class time on computer activities, discussion, in-class assessments (e.g., quizzes, tests, and/or clicker questions), and paper or other activities.

A flipped classroom is largely defined by instructors' use of class time to reinforce content that students encounter before the class meeting. Interestingly, students at two-year colleges in classrooms that were not flipped reported their instructors spent a similar amount of time lecturing (20 percent) as students in flipped classrooms (21 percent). From pre-pilot interviews and implementation tracking phone calls, we heard from instructors that some two-year students in classrooms that were not flipped were in classrooms using an emporium-style approach, in which in-class meetings were in computer classrooms or labs using adaptive software. In the emporium pedagogical approach, instructors deliver personalized on-on-one help to students with little or no large group discussions or lectures.

Finally, instructors and students reported on the timing of when students first engage with new statistics concepts, whether outside of class or in class. As summarized in table 8, student and instructor reports of the timing correlates with instructors' assessments of whether they offered a flipped classroom approach. This validates the pedagogy measure as representing meaningful differences in in-class experiences. Instructors at four-year institutions appear to have overestimated the incidence of student engagement with new statistics concepts outside of class for a flipped or partially flipped classroom approach.

TABLE 7

Spring 2018 Reported First Engagement with New Statistics Concepts Outside of Class

Two-year institutions

	Flipped	Partially flipped	Not flipped
Instructor-reported	75%	72%	46%
Student-reported	60%	49%	50%

Source: Instructor and Student End-of-Semester Survey Data

TABLE 8

Spring 2018 Reported First Engagement with New Statistics Concepts Outside of Class

Four-year institutions

	Flipped	Partially flipped	Not flipped
Instructor-reported	98%	95%	--
Student-reported	63%	56%	--

Source: Instructor and Student End-of-Semester Survey Data

Use of the Acrobatiq Tool

STUDENT TOOL USE

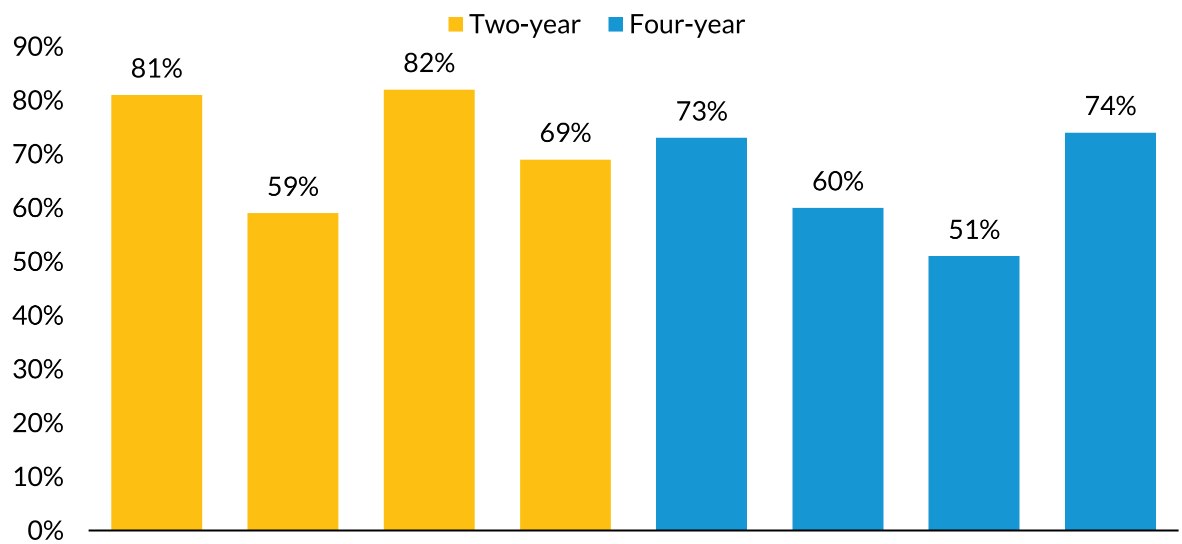
The ALiS project team set the expectation that pilot students would complete 90 percent of the content in the Acrobatiq tool over the course of the semester. As shown in figure 19, no college quite met that benchmark. The median percentage of activities completed was consistent across institution type, with 68 percent completed at two-year colleges and 67 percent at four-year colleges.

First attempt accuracy—the extent to which students got formative questions right on their first try—did not vary as much across institutions (63 percent at two-year colleges and 66 percent at four-year colleges).

FIGURE 19

Median Percent of Activities Completed, by Institution Type

Fall 2017 and spring 2018



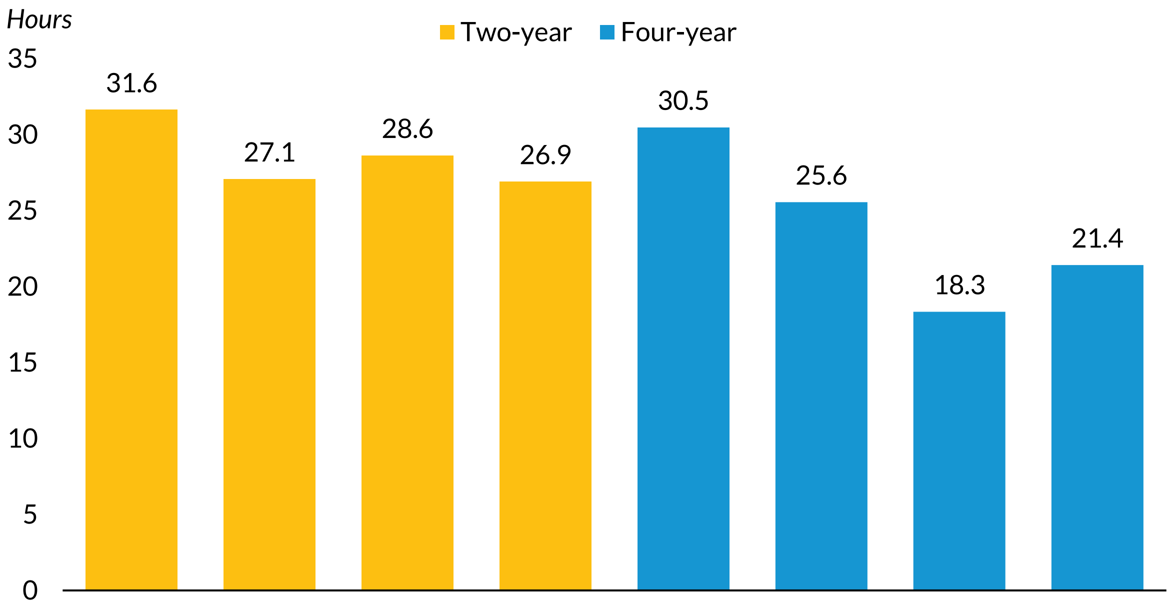
Note: Each column represents one institution.
Source: Acrobatiq Data

Figure 20 shows that two-year students spent more time in the tool overall (27 hours) than peer students in four-year institutions (23 hours).

FIGURE 20

Median Hours Spent in Platform, by Institution Type

Fall 2017 and spring 2018



Note: Each column represents one institution.
Source: Acrobatiq Data

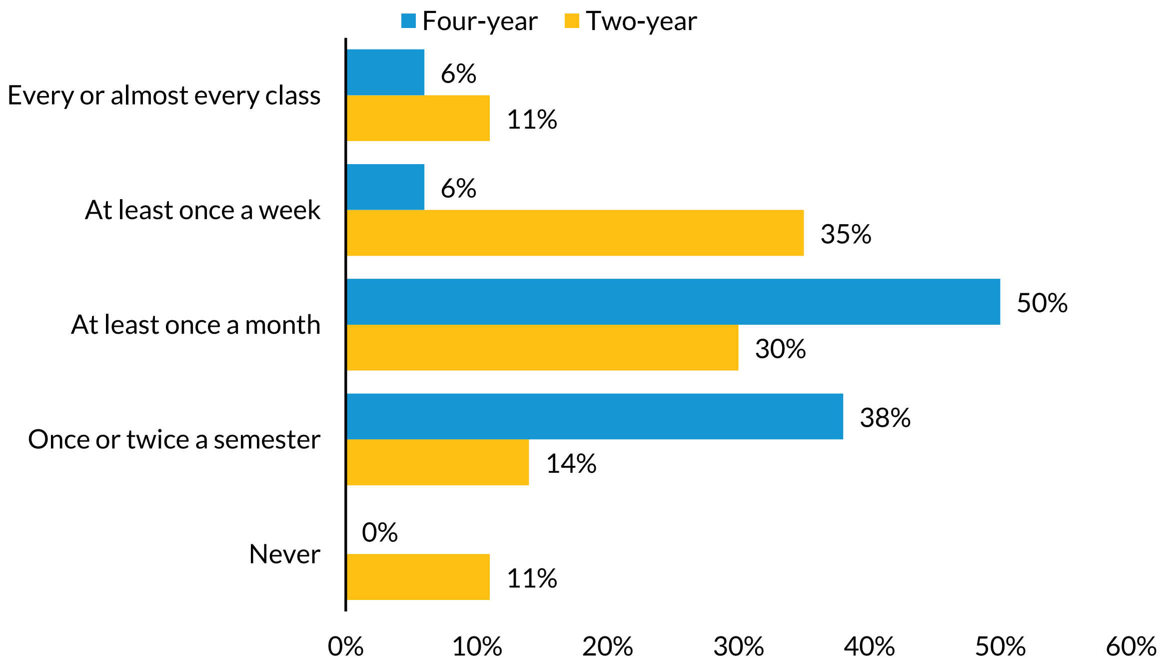
INSTRUCTOR TOOL USE

The ALiS model also called for instructors to use the Acrobatiq dashboard to track student progress, which would allow instructors to tailor in-class activities to student needs. Four-year instructors were more likely to use the dashboard at all; 100 percent of four-year instructors and 89 percent of two-year instructors reported using the dashboard to track student progress at some point during the semester. Contrarily, two-year instructors were more likely to use the tool with fidelity according to the theory of change for this project, using real-time student data on a regular basis to better understand students' progression through the course, as shown in figure 21. Two-year instructors used the Acrobatiq platform more frequently than four-year instructors, with 76 percent of two-year instructors reporting that they used the platform at least once a month, compared to 62 percent of four-year instructors. Roughly half of all two-year instructors (46 percent) reported using the platform weekly or every class period; compared with only 12 percent of four-year instructors who reported using the platform weekly or every class period.

FIGURE 21

Pilot Instructor Reported Use of the Dashboard to Track Student Progress

Fall 2017 and spring 2018



Source: Instructor Survey Data

Time and Effort

STUDENT-REPORTED TIME AND EFFORT

Across two- and four-year institutions, students reported more hours spent outside of classes that were not flipped or partially flipped, rather than flipped classes, as shown in Table 9. This is surprising, given that the fully flipped pedagogy would generally require students to spend more time becoming familiar with course concepts outside of class time.

TABLE 9

Student-Reported Weekly Hours Spent Outside of Class in the Acrobatiq Tool

Spring 2018, by institution type and pedagogy (medians)

	Flipped	Partially flipped	Not flipped
Two-year	6.9	7.3	7.9
Four-year	6.0	6.4	--

Source: Student end-of-semester data

As shown in Table 10, across both semesters and all pedagogical approaches, students at two-year institutions spent an average of 1.3 more hours a week completing work outside of class in the Acrobatiq tool than their four-year peers, while the percentage of activities completed in the Acrobatiq tool overall was roughly identical between two-year (68 percent) and four-year (67 percent) institutions. This additional time spent in the tool to complete roughly the same amount of activities across institution types may be due to factors such as reading comprehension differences. Two-year students may also have spent more time in the platform while they were in class, particularly those enrolled in emporium-style classes.

TABLE 10

Student-Reported Weekly Hours Spent Outside of Class in the Acrobatiq tool and Percent of Activities Completed

Fall 2017 and Spring 2018, by institution type (medians)

	Two-year	Four-year
Student-reported hours	7.5	6.4
Percent of activities completed	68%	67%

Source: Student end-of-semester data

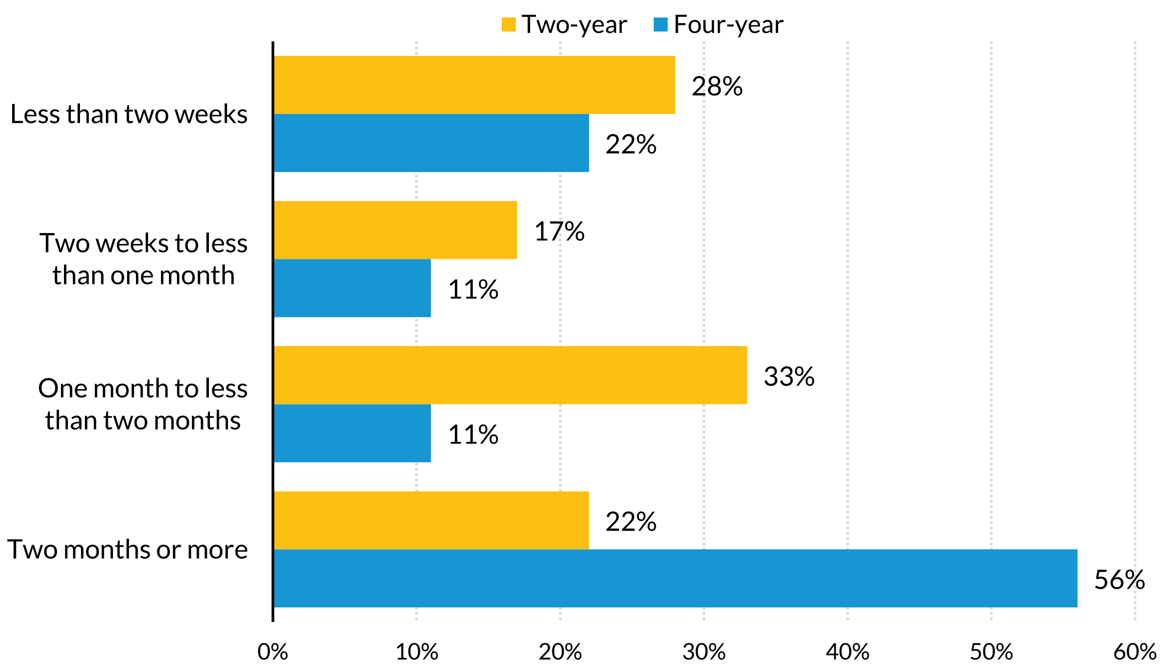
INSTRUCTOR PREPARATION TIME

Each semester, pilot instructors were asked to reflect on the amount of time they had to prepare for teaching the Acrobatiq course. The intervention involved adopting a new curriculum aligned with the Acrobatiq platform. For many instructors, the shift also required changing their course pedagogy to a flipped classroom approach. Figure 22 demonstrates that two-year instructors reported less preparation time overall. Only 22 percent of instructors at two-year colleges and 56 percent of four-year instructors reported having at least months to prepare to teach the course.

FIGURE 22

Instructor-Reported Prep Time with Acrobatiq, by Institution Type

Fall 2017 and spring 2018



Source: Instructor survey data

Feedback and Lessons

The open-response feedback on the fall 2017 and spring 2018 surveys provided additional context for the factors that may have affected student satisfaction in the pilot section of the course. The surveys asked students to provide any comments they wished to share about their course experience. Instructors were asked to describe in open-response feedback the elements that they found most promising and most challenging about their experience teaching the course. Common themes emerged in the feedback collected across both surveys.

READING COMPREHENSION AS A PREREQUISITE

Open-response feedback from the instructor surveys show that instructors may have felt compelled to offer more in-class instruction than they originally intended because students did not complete all the assigned reading and activities in the Acrobatiq tool prior to class. Particularly at two-year institutions, instructors shared that students' limited engagement with the platform made fully flipping the course difficult.

The need for high reading comprehension skills may have been difficult to overcome for students who started the course with lower reading comprehension skills than their classmates. One instructor said that the platform was more appropriate for students at more selective institutions (i.e., University of Maryland schools) or at the 300 level or above; rather than introductory statistics.

"The number one thing I recognized was that students were not up to par with their reading skills, and very quickly they just did not want to go on Acrobatiq and read the heavy number of pages they had to work through each day/week to stay on top of their work. For me it is the group of C and D students that I need to make stronger, and they normally need more time to improve reading and comprehension skills; and from my perspective, this was not accomplished with Acrobatiq." –Instructor at two-year college

"Students that have trouble reading may have struggled a bit with the way I used the software. I was really expecting them to [gain] understanding using the active learning [components]. Some tech phobic students had some difficulty, but they eventually got comfortable." –Instructor at two-year college

"I do not find that the ALiS approach is either more or less effective for any particular demographic. However, I do find that students with weaker reading skills have difficulty with the material and have a tendency to use (prefer) YouTube and Khan Academy content over the Acrobatiq text." –Instructor at four-year college

Students felt that the length of content in the Acrobatiq tool was a barrier to completing all necessary assignments. Specifically, some chapters in the online tool featured portions of text that were so time intensive that some opted to skip them. Students reported that some pages took 5 minutes while others took 30, making it difficult to gauge work length. Students had a difficult time keeping up the necessary pace to complete all of the assignments in Acrobatiq while also managing their other coursework. For nontraditional students, especially students working full-time and taking evening courses, Acrobatiq may not have been as effective of a learning tool due to insufficient time to digest assigned material before each class.

"The course is extremely fast-paced. Oftentimes we would be assigned 30+ pages of Acrobatiq a week not to mention the fact that I had other classes that needed my attention. As soon as I would begin to understand a topic, we would be on to the next, it was difficult for me to keep up with and I always felt a few steps behind." –Student at four-year college

"The work load is very heavy, especially with other classes and a full-time job." –Student at two-year college

SOFTWARE ADAPTIVITY AND TECHNICAL ISSUES

The details of the Acrobatiq tool are not directly relevant to the broader question of how adaptive online learning tools function, but specific challenges or successes with Acrobatiq might partially explain the satisfaction and other impact results in this study. One instructor described some of the adaptive features of the online platform, including the online data dashboard, as one of their favorite elements of the course.

"The main advantages of Acrobatiq for me have been the interactive features that require the students to pause frequently and do something with the knowledge being acquired, and the constantly adapting, objective-based dashboard data." –Instructor at two-year college

Multiple instructors were not convinced that there was as much value to Acrobatiq as an "adaptive" learning system and sought more adaptivity in the platform so that the product was differentiated from other, similar online learning tools for introductory statistics on the market. One two-year instructor simply noted, "Make the program truly adaptive." Another four-year instructor commented that they, "became more familiar with the platform, but was disappointed that there was not more adaptive content."

Students noticed typographical errors and omissions in the online textbook that were notable enough that they affected their ability to interpret concepts and complete activities.

"Overall, I think Acrobatiq could be useful, but it doesn't feel ready to be deployed. There were several times where sections on one page referred to examples on previous pages without appropriate references or links, so it was hard to know what was being discussed. In addition, it seemed as though different sections had different authors, as the flow and writing styles changed frequently. There were incorrect answers that made it confusing to learn. It was hard to flip back and forth between material. There were not enough opportunities to practice problems, and it was difficult to prepare for tests using the Acrobatiq platform, as there were not enough review opportunities." –Student at four-year college

STUDENT FRUSTRATION WITH INDEPENDENT LEARNING

The flipped classroom learning style and reliance on an online platform to introduce students to course concepts did not align with student expectations for the course, according to the majority of the open-response feedback. Students preferred to know that the course was going to be flipped prior to enrollment, and they craved lecture when concepts became difficult to learn on their own, especially toward the end of the semester. Many students spoke highly of their instructors and their engagement with the course content and subject, while still preferring a direct-instruction pedagogical style.

"I really enjoyed [my instructor's] class, but I did not enjoy the online supplemental material. I felt it took too much time and sometimes very confusing. I feel my grade would have been much better this semester if I had worked from a regular class with syllabus and study guides from the professor and text book. Not every system works for every person, as I am learning very quickly." –Student at four-year college

One returning instructor shared that students in his pilot section were experiencing a deeper understanding of the material because of the in-class exercises and flipped approach, but was not sure that the tool itself was instrumental to students' success as much as the emphasis on providing an active learning experience.

"I believe that the greatest asset of the ALiS course is that it provides an infrastructure for promoting active learning. It mandates independent work at home and some active processing and problem solving in class. The technology does not seem to me to be the determining factor in how successful a course is. The fact that there is so much support and direction toward providing an active learning experience, to me, seems to be at the core of its success. I believe that other courses (with or without such technologies) that provide integrated avenues for consistent, active engagement and processing could be just as successful." –Instructor at two-year college

Conclusion

This study adds to the body of research on adaptive learning approaches by highlighting the important roles of pedagogy, course curriculum, and course structure in the experience of students completing college-level, technology-mediated coursework. The modest findings across institution type and student subgroups are largely consistent with prior studies of blended learning approaches, in which student outcomes were not significantly adversely affected (nor significantly improved) due to participation in the course.

The intervention proved most promising for students in four-year colleges. The need to better understand the role of various elements of the course experience on student success and satisfaction should be investigated in future research. Student reading level, the flipped classroom course structure, and student motivation and ability to learn statistics concepts on their own outside of class are among the factors that could be impacting the effectiveness of the intervention.

Limitations and Directions for Future Research

Additional investigation is needed to discern the effects of the pedagogical approach and use of an online adaptive learning platform on student success. Future research could try to disaggregate the components of the intervention—the platform, the pedagogical approach, or the learning community—that have the largest effect on student outcomes. In addition, future research could identify the characteristics of students who most engage with, benefit from, and express satisfaction with this type of intervention.

Furthermore, while the pilot instructor survey in this study provides insight into the classroom experience of students completing ALiS sections of the course, we were not able to compare those findings to the experience of instructors teaching traditional sections of the course. Surveying instructors teaching traditional course sections would help deepen insights into instructor experiences. In addition, direct classroom observation of pilot and traditional sections, along with instructor interviews and student focus groups, would add depth to self-reported survey data on instructors' pedagogical approach and students' in-class experiences.

Finally, for many pilot instructors, the introduction of a new online course curriculum (Acrobatiq) and a new pedagogical approach (flipped classroom) meant that instructors were learning and adapting their course materials and course structure in real time. Future research could track instructors' approach to teaching the course over time to see if outcomes and impacts improve as instructors gain more experience. If short-term impacts improve, it would also be valuable to track students longer term to assess the impact this type of intervention can have in the long run on college graduation rates and transfer student outcomes, college completion rates, and future course taking.

FUTURE RESEARCH QUESTIONS

Several questions ripe for future research emerge from this study.

One area of future research relates to the most effective approaches for implementing adaptive learning approaches in the classroom and what types of institutions and students stand to benefit. Related research questions include:

- Is there still promise for technological innovations to improve student success and satisfaction?
- What would the next generation of successful interventions look like?
- For which institutions and students is the combination of a flipped classroom format supplemented with an adaptive online learning tool appropriate?
- What student supports could help promote nontraditional student success?

While all pilot instructors were given guidance on how to implement the Acrobatiq tool in their courses (see Figure 1), each instructor's fidelity to the suggested model of course delivery is difficult to assess without additional context. More in-depth qualitative research, including instructor and student interviews and classroom observations would have provided greater context to how the experience of a student in a pilot and traditional section of the course varies within and across institutions. This could help answer the following questions:

- What are the best pedagogical methods for integrating online adaptive learning platforms coursework into the curriculum?
- How can we better understand the factors (e.g., technological literacy, pedagogy, instructor and student characteristics) underlying who benefits and why so that outcomes are more equitable across institution type?

Furthermore, professional development resources provided to instructors for integrating adaptive learning tools into their curriculum should be explored in greater detail. For example, randomly assigning professional development resources to instructors implementing technology-mediated approaches in the classroom could show how professional development affects implementation of an unfamiliar online adaptive learning tool and a new model of instruction (i.e., the flipped classroom approach).

- How do we better prepare instructors to use the analytic dashboards to engage in formative assessment and continuous improvement of instruction?

While the evaluation gathered information about the learning resources that the ALiS Project Team and instructors made available to each other, the role of the professional learning community across institutions and the effect of peer learning were not key elements of the research. The ALiS pilot was unusual in the cross-institutional partnership that included two-year and four-year institutions and an emphasis from the ALiS project team peer-led support. It may be useful to understand more about the role of this type of collaborative structure in future research by investigating the following questions:

- Do professional learning communities across different types of institutions affect development and delivery of hybrid or technology-enhanced flipped courses?
- Which elements of peer learning structures are most effective in improving student course outcomes and instructor experiences?

Finally, the question of costs for colleges and students should be explored as a crucial element of sustainability in the long run. Future cost studies and cost-benefit analyses of similar approaches will help determine whether technology-enhanced flipped classroom instruction can be implemented in ways that are affordable for both institutions and students. Further studies can also help ascertain how cost or return-on-investment analysis in future research can help answer to this question.

Learning the answers to these questions would help build on the results in this study by providing additional context for understanding which technology-mediated interventions are most effective, for whom, and why. This research, along with future efforts, could have implications for the provision of math education and training at scale. Using technology to increase access to introductory college-level material for students—particularly those with nontraditional schedules and other commitments—can afford students flexibility, but there is more to learn about what additional supports may be needed for students to be successful and satisfied with their course-taking experience.

Appendix A. Pre-pilot Phase

The additional information collected via site visit observations, interviews, and focus groups helped the study team better understand the program intervention and refine our understanding of the theory of change for the fall 2017 and spring 2018 pilot year.

Data Sources

The research team collected data from two institutions—University of Maryland College Park and Montgomery College—in the pre-pilot phase. Table 11 summarizes data collection activities.

TABLE 11

Qualitative Data Collection Activities Completed Fall 2016 and Spring 2017

Instrument or Source	Population	Format
Classroom observations	Students and faculty in multiple traditional and pilot sections at each institution	Observation time that corresponded with class duration
Focus groups	One with students in traditional sections and one in pilot sections in each institution	One hour
Individual and small-group interviews	Instructors of traditional and pilot sections and administrators	One hour

Interview Topic Guides

Student Focus Group Topic Guide

- Background in statistics
- General impressions of the course
- Experience using Acrobatiq, MyStatLab, or other online homework tool
- Benefits and disadvantages of using the tool in the classroom
- Perceived impact of the tool on learning
- Desire to participate in other classes using a similar model of instruction

- Future plans for math education, college, and career
- Ways to improve classroom experience for future cohorts

Faculty Interview Topic Guide

- Experience teaching course (impact on pedagogy)
- General impressions of tool (MyStatLab/StatKey/Acrobatiq)
- Benefits and disadvantages of using tool
- Perceived advantages and disadvantages in student experience and learning outcomes
- Opportunities for and challenges to scalability
- Lessons learned for the future

Administrator Interview Topic Guide

- Role at institution and responsibility overseeing pilot
- Motivation for introducing Acrobatiq into curriculum and historical experience with piloting online or hybrid courses in math departments
- Goals and expectations for the pilot
- Implementation approaches by participating instructors and how this has evolved over time (if returning institution)
- Successes and challenges with implementation of Acrobatiq
- Instructor resources available
- Cost of implementing Acrobatiq compared to traditional statistics curriculum (i.e. dollars and staff time)

In-Class Observation Topic Guide

- Activity identification, including general nature of lecture and lab
- Logistical features, including materials used for session, timeliness, and number of students

- Form and content, including how topics were covered and instruction was delivered, goals for activities, and assigned work
- Student engagement, including student attitudes and communication methods
- Reflections on suggested modifications and other observations

Key Findings

Some of our observations and key findings from the qualitative research and data collection conducted in the pre-pilot phase included the following, which helped to inform implementation adjustments made in the fall 2017 to spring 2018 project period:

- It would be helpful if instructors familiarized themselves with the modules in Acrobatiq in advance, to (1) anticipate technical difficulties, as discussed above;(2) direct students' attention to the key learning objectives for each module; and (3) tailor their lecture materials to align better with the Acrobatiq content.
- Acrobatiq may want to provide instructors with more prepackaged material, such as slides and in-class reinforcement exercises, to enhance the usability of the tool in lecture and the ability for instructors to navigate through the content. Other publishers provide these types of materials, which many instructors report are very useful.
- Instructors may want to consider redesigning courses so that most of the content is guided by Acrobatiq, and then adapt course materials and experiences to the Acrobatiq tool. In this model, "Acrobatiq is the GPS and the instructor is driving the car," meaning Acrobatiq guides the course and the instructor carries it out, but with appropriate corrections as needed. This may require a fundamental shift in the course ordering, content, and may only be appropriate once some of the alignment challenges (e.g., the theoretical versus applied approach) are overcome.
- Implementers of the tool and the evaluation team will need to remain cognizant of the interaction between the use of the tool and course structure (instructor–student ratio, roles of lab and lecture, meeting time, etc.) in affecting students' experiences and outcomes. In particular, the quality of the evaluation will be undermined if there are substantive differences besides usage of the tool between the treatment and comparison classrooms.
- The instructors may want to better align their lectures and in-class activities with the Acrobatiq content. A small change would be to match the labels on the content sections of the lectures and

the tool to help students more readily find content they are searching for in the platform. A larger change would be more purposeful integration of the tool into the course structure or a revision of the course to align better with the tool.

Appendix B. Impact Study Design and Additional Findings

Covariates

Demographics:

- Age (continuous)
- Female (0/1)
- White, non-Hispanic (0/1)
- Asian, non-Hispanic (0/1)
- Other/multiple race, non-Hispanic (0/1)
- Hispanic of any race (0/1)
- Black, non-Hispanic is the omitted reference group
- Born in the US (0/1)
- First language is English (0/1)

Subgroup status:

- First generation college student (0/1)
- Pell eligible (0/1)
- Prior developmental math experience (0/1)

Prior statistics experience:

- Took statistics in high school or college (0/1)

Prior college performance:

- Number of credits earned before the start of the semester (continuous)
- Cumulative GPA (continuous, 0.0–4.0 scale)

Standardized test scores:

- Accuplacer elementary algebra score (continuous)
- SAT/ACT percentile (continuous, 0–100)

Employment, full-time student status:

- Not expecting to work in the semester at baseline (0/1)
- Expecting to work less than 20 hours per week in the semester at baseline (0/1)
- Expecting to work 20 hours or more per week in the semester at baseline is the omitted reference group

- Full time student (0/1)

Baseline aptitude, self-efficacy, attitudes:

- Baseline aptitude (continuous, 0-6 scale)
- Baseline attitude (continuous, 0-5 scale)
- Baseline confidence (continuous, 0-5 scale)

Other

- Indicator for each matched pair (0/1)

We also included indicator variables for missing values and set the main variables to 0 (for binary characteristics) or the mean (for continuous characteristics).

Detailed Coefficients and Standard Errors

Primary Results, Fall Semester

	Fall Semester															
	Grade (GPA on 4.0 scale)				Pass C or Better (probability)				Stats Competency (standardized)				Satisfaction (five-point scale)			
	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N
All Colleges	0.130	(0.071)	*	1,465	0.026	(0.025)		1,465	0.067	(0.043)		1,287	-0.033	(0.077)		1,141
All Colleges, within Same-Instructor Matched Pairs	-0.049	(0.110)		488	-0.053	(0.040)		488	0.154	(0.072)	**	397	-0.600	(0.120)	***	356
All Colleges except Towson and Wor-Wic	0.130	(0.071)	*	1,465	0.026	(0.025)		1,465	0.067	(0.043)		1,287	-0.033	(0.077)		1,141
Two-Year Colleges	0.011	(0.107)		611	-0.017	(0.038)		611	-0.031	(0.068)		493	-0.495	(0.119)	***	431
Four-Year Colleges	0.291	(0.092)	***	854	0.078	(0.032)	***	854	0.173	(0.054)	***	794	0.440	(0.098)	***	710

Primary Results, Spring Semester

	Spring Semester															
	Grade (GPA on 4.0 scale)				Pass C or Better (probability)				Stats Competency (standardized)				Satisfaction (five-point scale)			
	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N
All Colleges	0.033	(0.057)		2,210	0.004	(0.021)		2,210	0.087	(0.034)	***	1,944	-0.191	(0.063)	***	1,730
All Colleges, within Same-Instructor Matched Pairs	-0.118	(0.072)	*	1,214	-0.060	(0.028)	**	1,214	0.043	(0.049)		983	-0.348	(0.084)	***	862
All Colleges except Towson and Wor-Wic	0.122	(0.068)	*	1,619	0.043	(0.025)	*	1,619	0.091	(0.040)	**	1,474	-0.124	(0.073)	*	1,330
Two-Year Colleges	-0.080	(0.086)		1,038	-0.029	(0.033)		1,038	-0.005	(0.056)		815	-0.316	(0.098)	***	720
Four-Year Colleges	0.176	(0.071)	***	1,172	0.049	(0.025)	*	1,172	0.178	(0.040)	***	1,129	-0.013	(0.079)		1,010

Primary Results, Both Semesters

	Both Semesters															
	Grade (GPA on 4.0 scale)				Pass C or Better (probability)				Stats Competency (standardized)				Satisfaction (five-point scale)			
	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N
All Colleges	0.081	(0.044)	*	3,675	0.017	(0.016)		3,675	0.075	(0.027)	***	3,231	-0.093	(0.049)	*	2,871
All Colleges, within Same-Instructor Matched Pairs	-0.101	(0.060)	*	1,702	-0.060	(0.022)	***	1,702	0.089	(0.040)	**	1,380	-0.434	(0.069)	***	1,218
All Colleges except Towson and Wor-Wic	0.132	(0.049)	***	3,084	0.038	(0.017)	**	3,084	0.076	(0.029)	***	2,761	-0.050	(0.052)		2,471
Two-Year Colleges	-0.038	(0.067)		1,649	-0.022	(0.024)		1,649	-0.020	(0.042)		1,308	-0.355	(0.074)	***	1,151
Four-Year Colleges	0.161	(0.050)	***	2,026	0.038	(0.018)	**	2,026	0.112	(0.031)	***	1,923	0.264	(0.055)	***	1,720

Subgroup Results, Fall Semester

Note that we exclude significance indicators because significance is primarily reported between groups as opposed to being reported for each individual measure.

	Fall Semester											
	Grade (GPA on 4.0 scale)			Pass C or Better (probability)			Stats Competency (standardized)			Satisfaction (five-point scale)		
	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N
ALL COLLEGES												
Not First Gen	0.127	(0.079)	1,210	0.024	(0.027)	1,210	0.107	(0.048)	1,073	-0.039	(0.084)	959
First Gen	0.058	(0.172)	255	0.032	(0.065)	255	-0.138	(0.108)	214	0.213	(0.212)	182
Not Pell	0.130	(0.079)	1,168	0.029	(0.027)	1,168	0.130	(0.047)	1,050	0.049	(0.083)	931
Pell	0.212	(0.151)	297	0.058	(0.057)	297	-0.043	(0.113)	237	-0.253	(0.231)	210
Not Dev Ed	0.027	(0.094)	907	0.019	(0.031)	907	0.065	(0.056)	833	0.174	(0.097)	740
Dev Ed	0.212	(0.111)	558	0.028	(0.042)	558	0.040	(0.070)	454	-0.189	(0.136)	401
New Instructor	0.209	(0.092)	829	0.064	(0.032)	829	0.109	(0.057)	703	-0.355	(0.103)	612
Returning Instructor	-0.097	(0.100)	636	-0.077	(0.036)	636	-0.040	(0.063)	584	0.738	(0.100)	529

	Fall Semester											
	Grade (GPA on 4.0 scale)			Pass C or Better (probability)			Stats Competency (standardized)			Satisfaction (five-point scale)		
	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N
TWO-YEAR COLLEGES												
Not First Gen	0.015	(0.125)	454	-0.027	(0.044)	454	0.051	(0.082)	372	-0.485	(0.139)	329
First Gen	-0.121	(0.244)	157	0.014	(0.091)	157	-0.368	(0.160)	121	-0.233	(0.284)	102
Not Pell	0.033	(0.136)	377	-0.002	(0.046)	377	0.001	(0.087)	312	-0.454	(0.152)	275
Pell	0.042	(0.184)	234	-0.010	(0.070)	234	-0.055	(0.138)	181	-0.580	(0.262)	156
Not Dev Ed	-0.181	(0.214)	211	-0.028	(0.068)	211	-0.195	(0.133)	184	-0.516	(0.227)	168
Dev Ed	0.066	(0.133)	400	-0.022	(0.050)	400	0.020	(0.089)	309	-0.381	(0.164)	263
New Instructor	0.055	(0.126)	472	0.011	(0.044)	472	-0.037	(0.083)	379	-0.826	(0.141)	329
Returning Instructor	-0.227	(0.240)	139	-0.130	(0.084)	139	-0.064	(0.153)	114	0.501	(0.259)	102
FOUR-YEAR COLLEGES												
Not First Gen	0.286	(0.098)	756	0.088	(0.034)	756	0.162	(0.059)	701	0.411	(0.102)	630
First Gen	0.076	(0.259)	98	-0.057	(0.097)	98	0.228	(0.190)	93	0.299	(0.360)	80
Not Pell	0.188	(0.097)	791	0.047	(0.033)	791	0.211	(0.057)	738	0.453	(0.096)	656
Pell	0.936	(0.303)	63	0.243	(0.119)	63	0.011	(0.223)	56	0.437	(0.494)	54
Not Dev Ed	0.125	(0.103)	696	0.034	(0.035)	696	0.178	(0.062)	649	0.557	(0.104)	572
Dev Ed	0.648	(0.199)	158	0.177	(0.079)	158	0.135	(0.118)	145	0.236	(0.234)	138
New Instructor	0.449	(0.141)	357	0.135	(0.049)	357	0.270	(0.081)	324	0.182	(0.148)	283
Returning Instructor	-0.051	(0.088)	497	-0.059	(0.032)	497	-0.037	(0.062)	470	0.913	(0.100)	427

Subgroup Results, Spring Semester

Note that we exclude significance indicators because significance is primarily reported between groups as opposed to being reported for each individual measure.

	Spring Semester											
	Grade (GPA on 4.0 scale)			Pass C or Better (probability)			Stats Competency (standardized)			Satisfaction (five-point scale)		
	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N
ALL COLLEGES												
Not First Gen	0.033	(0.063)	1,810	0.015	(0.024)	1,810	0.109	(0.037)	1,620	-0.136	(0.071)	1,431
First Gen	0.032	(0.137)	400	-0.026	(0.052)	400	0.002	(0.085)	324	-0.366	(0.163)	299
Not Pell	0.074	(0.076)	1,337	0.027	(0.027)	1,337	0.098	(0.044)	1,199	-0.127	(0.087)	1,032

	Spring Semester											
	Grade (GPA on 4.0 scale)			Pass C or Better (probability)			Stats Competency (standardized)			Satisfaction (five-point scale)		
	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N
Pell	-0.034	(0.082)	873	-0.033	(0.033)	873	0.076	(0.054)	745	-0.289	(0.095)	698
Not Dev Ed	0.158	(0.072)	1,394	0.032	(0.025)	1,394	0.147	(0.041)	1,288	-0.037	(0.075)	1,134
Dev Ed	-0.105	(0.094)	816	-0.026	(0.038)	816	0.027	(0.061)	656	-0.395	(0.111)	596
New Instructor	-0.158	(0.092)	688	-0.078	(0.037)	688	0.051	(0.060)	537	-0.474	(0.124)	455
Returning Instructor	0.110	(0.070)	1,522	0.035	(0.026)	1,522	0.102	(0.041)	1,407	-0.114	(0.074)	1,275
TWO-YEAR COLLEGES												
Not First Gen	-0.131	(0.101)	770	-0.021	(0.038)	770	-0.026	(0.064)	616	-0.235	(0.118)	538
First Gen	-0.023	(0.175)	268	-0.073	(0.068)	268	0.085	(0.120)	199	-0.447	(0.227)	182
Not Pell	-0.145	(0.127)	496	-0.044	(0.047)	496	-0.046	(0.081)	399	-0.223	(0.152)	336
Pell	-0.055	(0.108)	542	-0.027	(0.044)	542	0.034	(0.077)	416	-0.484	(0.136)	384
Not Dev Ed	-0.061	(0.147)	381	-0.040	(0.049)	381	-0.018	(0.091)	300	-0.287	(0.162)	250
Dev Ed	-0.067	(0.109)	657	-0.020	(0.044)	657	0.032	(0.073)	515	-0.349	(0.125)	470
New Instructor	-0.083	(0.147)	346	-0.043	(0.058)	346	-0.121	(0.100)	247	-0.700	(0.217)	207
Returning Instructor	-0.045	(0.105)	692	-0.012	(0.040)	692	0.027	(0.069)	568	-0.217	(0.112)	513
FOUR-YEAR COLLEGES												
Not First Gen	0.202	(0.075)	1,040	0.062	(0.027)	1,040	0.204	(0.042)	1,004	-0.017	(0.084)	893
First Gen	0.152	(0.227)	132	0.063	(0.093)	132	-0.093	(0.134)	125	0.126	(0.296)	117
Not Pell	0.249	(0.087)	841	0.092	(0.030)	841	0.168	(0.048)	800	-0.021	(0.099)	696
Pell	-0.063	(0.121)	331	-0.062	(0.048)	331	0.138	(0.077)	329	0.049	(0.145)	314
Not Dev Ed	0.264	(0.077)	1,013	0.071	(0.027)	1,013	0.202	(0.044)	988	0.092	(0.083)	884
Dev Ed	-0.268	(0.196)	159	-0.043	(0.080)	159	-0.018	(0.144)	141	-0.608	(0.260)	126
New Instructor	-0.153	(0.123)	342	-0.075	(0.051)	342	0.191	(0.076)	290	-0.215	(0.159)	248
Returning Instructor	0.317	(0.086)	830	0.102	(0.030)	830	0.170	(0.049)	839	0.063	(0.092)	762

Subgroup Results, Both Semesters

Note that we exclude significance indicators because significance is primarily reported between groups as opposed to being reported for each individual measure.

	Both Semesters											
	Grade (GPA on 4.0 scale)			Pass C or Better (probability)			Stats Competency (standardized)			Satisfaction (five-point scale)		
	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N	Estimate	SE	N
ALL COLLEGES												
Not First Gen	0.077	(0.049)	3,020	0.023	(0.018)	3,020	0.106	(0.030)	2,693	-0.077	(0.053)	2,390
First Gen	0.073	(0.108)	655	-0.003	(0.040)	655	-0.057	(0.063)	538	-0.075	(0.127)	481
Not Pell	0.114	(0.055)	2,505	0.034	(0.019)	2,505	0.125	(0.032)	2,249	-0.028	(0.060)	1,963
Pell	0.017	(0.074)	1,170	-0.018	(0.029)	1,170	0.004	(0.049)	982	-0.227	(0.087)	908
Not Dev Ed	0.114	(0.056)	2,301	0.027	(0.019)	2,301	0.112	(0.034)	2,121	0.067	(0.059)	1,874
Dev Ed	0.064	(0.072)	1,374	0.012	(0.028)	1,374	0.029	(0.045)	1,110	-0.302	(0.085)	997
New Instructor	0.118	(0.070)	1,517	0.023	(0.025)	1,517	0.061	(0.045)	1,240	-0.345	(0.083)	1,067
Returning Instructor	0.086	(0.058)	2,158	0.018	(0.021)	2,158	0.090	(0.034)	1,991	0.085	(0.061)	1,804
TWO-YEAR COLLEGES												
Not First Gen	-0.087	(0.077)	1,224	-0.028	(0.028)	1,224	0.007	(0.050)	988	-0.337	(0.087)	867
First Gen	-0.016	(0.140)	425	-0.016	(0.051)	425	-0.126	(0.086)	320	-0.370	(0.174)	284
Not Pell	-0.023	(0.093)	873	-0.012	(0.032)	873	0.010	(0.057)	711	-0.345	(0.107)	611
Pell	-0.090	(0.096)	776	-0.048	(0.037)	776	-0.043	(0.066)	597	-0.435	(0.111)	540
Not Dev Ed	-0.087	(0.111)	592	-0.032	(0.037)	592	-0.097	(0.075)	484	-0.355	(0.128)	418
Dev Ed	0.003	(0.084)	1,057	-0.012	(0.032)	1,057	0.017	(0.053)	824	-0.336	(0.097)	733
New Instructor	0.047	(0.101)	818	-0.006	(0.035)	818	-0.097	(0.069)	626	-0.739	(0.121)	536
Returning Instructor	-0.049	(0.090)	831	-0.019	(0.033)	831	0.046	(0.056)	682	-0.088	(0.096)	615
FOUR-YEAR COLLEGES												
Not First Gen	0.252	(0.060)	1,796	0.080	(0.021)	1,796	0.197	(0.035)	1,705	0.161	(0.064)	1,523
First Gen	0.096	(0.162)	230	-0.026	(0.062)	230	0.113	(0.102)	218	0.353	(0.211)	197
Not Pell	0.225	(0.064)	1,632	0.074	(0.022)	1,632	0.207	(0.037)	1,538	0.208	(0.068)	1,352
Pell	0.190	(0.115)	394	0.028	(0.046)	394	0.102	(0.075)	385	0.159	(0.143)	368
Not Dev Ed	0.213	(0.061)	1,709	0.060	(0.021)	1,709	0.202	(0.036)	1,637	0.276	(0.064)	1,456
Dev Ed	0.301	(0.147)	317	0.089	(0.056)	317	0.117	(0.087)	286	-0.088	(0.164)	264
New Instructor	0.239	(0.097)	699	0.065	(0.035)	699	0.232	(0.058)	614	0.054	(0.111)	531
Returning Instructor	0.230	(0.068)	1,327	0.061	(0.023)	1,327	0.120	(0.040)	1,309	0.286	(0.073)	1,189

Satisfaction Results, Both Semesters

	Both Semesters															
	Satisfaction with Tool				Quality of Instruction				Instructor Availability				Activities Were Engaging			
	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N
All Colleges	-0.426	(0.064)	***	2,610	0.123	(0.042)	***	2,910	0.221	(0.042)	***	2,875	0.098	(0.045)	**	2,920
Two-Year Colleges	-0.675	(0.103)	***	1,036	-0.059	(0.064)		1,174	-0.025	(0.063)		1,169	-0.154	(0.069)	**	1,178
Four-Year Colleges	-0.139	(0.075)	*	1,574	0.304	(0.053)	***	1,736	0.472	(0.056)	***	1,706	0.358	(0.057)	***	1,742

	Both Semesters															
	Activities Helped Comprehension				Amount Learned				Difficulty of Course				Interest in Subject			
	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N	Estimate	SE	Sig	N
All Colleges	0.048	(0.049)		2,792	-0.183	(0.043)	***	2,912	0.033	(0.036)		2,926	-0.099	(0.046)	**	2,926
Two-Year Colleges	-0.126	(0.075)	*	1,144	-0.355	(0.068)	***	1,172	0.103	(0.056)	*	1,179	-0.216	(0.073)	***	1,178
Four-Year Colleges	0.222	(0.063)	***	1,648	-0.015	(0.053)		1,740	-0.045	(0.047)		1,747	0.033	(0.059)		1,748

	Both Semesters			
	Future Math Classes			
	Estimate	SE	Sig	N
All Colleges	-0.090	(0.052)	*	2,908
Two-Year Colleges	-0.126	(0.081)		1,173
Four-Year Colleges	-0.029	(0.067)		1,735

References

Bernard, Robert M., Philip Abrami, Eugene Borokhovski, C. Anne Wade, Rana M. Tamim, Michael A. Surkes, Edward Clement Bethel. 2009. "A Meta-Analysis of Three Types of Interaction Treatments in Distance Education." *Review of Educational Research* 79(3): 1243–89.

Bowen, G., M. Chingos, Lack, K., andNygren, T. "*Interactive Learning Online at Public Universities: Evidence from Randomized Trials*," Ithaka S+R, New York, 2012, available at https://sr.ithaka.org/wp-content/uploads/2015/08/sr-ithaka-interactive-learning-online-at-public-universities.pdf

Brame, C., (2013). Flipping the classroom. Vanderbilt University Center for Teaching. Retrieved November 2018 from http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/

Chen, Kuo-Su, Lynn Monrouxe, Yi-Hsuan Lu, Chang-Chyi Jenq, Yeu-Jhy Chang, Yu-Che Chang, and Pony Yee-Chee Chai. 2018). "Academic Outcomes of Flipped Classroom Learning: A Meta-Analysis." *Medical Education* 52(9): 910–24.

Chingos, Matthew, Rebecca Griffiths, Christine Mulhern, and Richard R. Spies. 2016. "Interactive Online Learning on Campus: Comparing Students' Outcomes in Hybrid and Traditional Courses in the University System of Maryland." *The Journal of Higher Education* 88(2): 210–33.

Chingos, Matthew, Rebecca J. Griffiths, Christine Mulhern, and Richard R. Spies. 2014. *Interactive Online Learning on Campus: Testing MOOCs and Other Platforms in Hybrid Formats in the University System of Maryland*." New York: Ithaka S+R. https://sr.ithaka.org/publications/interactive-online-learning-on-campus/

Nielsen, Doreen and Yazrina Yahya. 2013. "Co-creating and Mapping Curricula to the VLE." *Procedia Technology* 11: 710–17.

Garnham, Carla, and Robert Kaleta. 2002. "Introduction to Hybrid Courses." *Teaching With Technology Today* 8 (6) (March 20). http://www.wisconsin.edu/ttt/articles/garnham.htm.

Goodman, Joshua, Julia Melkers, and Amanda Pallais. 2018. Can Online Delivery Increase Access to Education? *Journal of Labor Economics* 37(1): 1–34.

Griffiths, Chingos, and C. Mulhern (2015). "*Randomized Trials Using Pearson's MyFoundationsLab in Summer Bridge Programs*" Ithaka S+R, New York, available at http://www.sr.ithaka.org/publications/can-online-learning-improve-college-math-readiness/

Johnson, Hans and Marisol Cuellar Mejia. 2014. *Online Learning and Student Outcomes in California 's Community Colleges.* San Fransisco, CA: Public Policy Institute of California.

Limniou, Maria, Ian Schermbrucker. and Minna Lyons. 2018. "Traditional and Flipped Classroom Approaches Delivered by Two Different Teachers: The Student Perspective." *Education and Information Technologies* 23(2):797–817.

Margulieux, L. E., McCracken, W. M., & Catrambone, R. (2015). Mixing in class and online learning: Content meta-analysis of outcomes for hybrid, blended, and flipped courses. In *Proceedings of Computer Supported Collaborative Learning*. International Society of the Learning Sciences.

Potter, Jodi. 2015. "Applying a Hybrid Model: Can it Enhance Student Learning Outcomes?" *Journal of Instructional Pedagogies* 17.

Protopsaltis, Spiros and Sandy Baum 2019. *Does Online Education Live Up to its Promise? A Look at the Evidence and Implications for Federal Policy*. Fairfax, VA: George Mason University.

University of California. 2016. "Adaptive Learning Technology Pilot Report." https://www.ucop.edu/institutional-research-academic-planning/_files/BFI-Adaptive-Learning-Technology-Report.pdf.

Xu, Di and Shanna Smith Jaggars. 2011. "Online and Hybrid Course Enrollment and Performance in Washington State Community and Technical College." Working Paper No. 31. New York, NY: Community College Research Center, Teachers College, Columbia University.

Yarnall, Louise, Barbara Means, and Tallie Wetzel. "Lessons Learned from Early Implementations of Adaptive Courseware." SRI Education. April, 2016. https://www.sri.com/sites/default/files/brochures/almap_final_report.pdf.

Notes

[1] Note that Acrobatiq was merged into VitalSource Technologies in September 2018, https://www.insidehighered.com/digital-learning/article/2018/09/05/vitalsource-acquires-courseware-platform-acrobatiq.

[2] A report prepared by the ALiS project team describing the development and implementation of the project along with their findings is available at https://sr.ithaka.org/publications/adaptive-learning-courseware-introductory-statistics. The findings presented in this report are focused on the impact and implementation evaluation conducted by the Urban Institute during the pilot year 2017-2018.

[3] Ibid.

[4] Ithaka S+R partnered with Transforming Post-Secondary Education in Math (TPSE Math), the William E. Kirwan Center for Academic Innovation at the University System of Maryland (USM), the University of Maryland, College Park (UMCP), Montgomery College (MC), and Acrobatiq to implement the ALiS Project. This group is described throughout the report as the "ALiS Project Team".

[5] See definition #5, "Definitions of E-Learning Courses and Programs Version 2.0." https://onlinelearningconsortium.org/updated-e-learning-definitions-2/

[6] For more information on Acrobatiq and the online platform, see "Solutions for Optimizing Teaching, Learning and Student Achievement," accessed June 10, 2019. http://acrobatiq.com/wp-content/uploads/2015/11/About-Acrobatiq.pdf

[7] A significant amount of the course content in the pilot sections was supposed to be delivered online by the Acrobatiq platform, consistent with a hybrid learning modality (Garnham and Kaleta 2002). However, only a small number of ALiS sections were deliberately structured as "hybrid" courses, meeting they only met face-to-face between 25% and 75% of the scheduled class time.

[8] The ALiS course had elements of both blended and hybrid learning, The term "hybrid/blended" in this report is intended to capture elements of blended and hybrid course delivery approaches. See "Blended Learning, Hybrid Learning, The Flipped Classroom… What's the Difference?" *Panopto* (blog), April 7, 2017, https://www.panopto.com/blog/blended-learning-hybrid-learning-flipped-classroom-whats-difference/.

[9] The "Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016" informed the development of the ALiS model, in terms of the content and approach for teaching introductory college level statistics in the course. See report linked **here**.

[10] DCMP was developed by the Charles A. Dana Center at the University of Texas at Austin and the Carnegie Foundation for the Advancement of Teaching.[10] In the Dana Center Math Pathways (DCMP)model, all students enter a credit-bearing mathematics course that prioritizes active learning and the relevant application of course concepts.

[11] Ibid.

[12] "First in the World Maryland Mathematics Reform Initiative," The University System of Maryland, accessed July 2019, https://dcmathpathways.org/sites/default/files/2016-08/First%20in%20the%20World%20Maryland%20Mathematics%20Reform%20Initiative%20%28FITW%20MMRI%29%20Project%20Overview.pdf

[13] The Urban Institute continued the evaluation research at five colleges in fall 2018, but the methods and research design differed substantially, so those results are not included in this report. See slide deck linked **here** for more information.

[14] Students' statistics competency was measured using a standardized assessment of 15 common questions administered across all sections. Instructors incorporated the questions, which were worth roughly half of the overall final scores, into the final exams. The questions were developed from standardized assessment question test banks, including the LOCUS, 2016 GAISE Appendix, 2012 AP Exam, and input from statistics instructors at participating institutions. We used item response theory (IRT) modeling to validate the questions and answer choices each semester. IRT assesses the weight of questions in predicting student ability as a product of discrimination and difficulty, and we used students' IRT-weighted scores as an indicator of statistics competency. We also used results of the IRT to refine the questions and answer choices between the fall and spring semester.

[15] In traditional sections, first-generation and Pell-eligible students, and those with prior developmental math experience had significantly worse grades, pass rates, and statistics competency in the aggregate. The patterns were consistent at both two-year and four-year colleges. Larger differences appeared among Pell-eligible and students with prior developmental education experience relative to their more advantaged peers. Smaller—but still significant—differences were observed among the first-generation student subgroup. Student satisfaction among these subgroups did not show a similarly consistent pattern of difference.

[16] We are grateful to Matt Griffin for contributing much of the content for the math self-efficacy and math attitude scales used in the surveys.

[17] LOCUS, accessed September 3,2019, https://locus.statisticseducation.org/

[18] The error term is not clustered because the model controls for each matched pair.

[19] Pedagogy questions were added in spring 2018.

[20] Wor-Wic Community College and Towson University joined the study in spring 2018.

[21] Note that many traditional sections also used a web-based homework tool, though not an adaptive textbook replacement.

[22] The confidence and attitude measures were provided by Dr. Matthew Griffin at the University of Maryland College Park.

About the Authors

Amanda Briggs is a research associate in the Income and Benefits Policy Center at the Urban Institute. Her research focuses on workforce development policy analysis and program evaluation, employer involvement in workforce training, and online learning. She currently is directing a study investigating the different types of supports provided to community college students completing coursework online, the role of technology in providing supports, and the evidence of effectiveness. Briggs earned a master's in public affairs with a concentration in social and economic policy from the Lyndon B. Johnson School of Public Affairs at the University of Texas at Austin.

Theresa Anderson is a senior research associate at the Urban Institute. She is a member of the Building America's Workforce cross-center initiative and is affiliated with the Income and Benefits Policy Center and the Education Data and Policy Center. She conducts in-depth, mixed-methods research on important policy issues, primarily evaluations of workforce, education, and social safety net programs and policies. Concurrent with her work at Urban, Anderson is pursuing a PhD in public policy and public administration at George Washington University, where she earned her MPP. She also received a BA from Hampshire College in Massachusetts.

Semhar Gebrekristos is a research analyst in the Income and Benefits Policy Center at the Urban Institute. She conducts quantitative and qualitative research on such topics as career pathways and the intersection of workforce development and child care. Before joining Urban, she was a research assistant in the education program at the American Institutes for Research. She is experienced in statistical testing and reporting and holds a bachelor's degree in economics from Mount Holyoke College.

Alphonse Simon is a research assistant with extensive experience conducting quantitative analysis. Simon has experience with API data pulls, web scraping, regression analysis, and developing data visualization tools. Previously, he conducted industrial organization research as a Research Associate at Columbia Business School. Simon graduated with a BA in Economics from the University of California, Berkeley.

Alice Mei was a research assistant who contributed to the Evaluation of Adaptive Learning in Statistics project during the fall 2017-spring 2018 school year. Mei holds a BA in Political Science and International Relations and Psychology from Carleton College.

STATEMENT OF INDEPENDENCE

The Urban Institute strives to meet the highest standards of integrity and quality in its research and analyses and in the evidence-based policy recommendations offered by its researchers and experts. We believe that operating consistent with the values of independence, rigor, and transparency is essential to maintaining those standards. As an organization, the Urban Institute does not take positions on issues, but it does empower and support its experts in sharing their own evidence-based views and policy recommendations that have been shaped by scholarship. Funders do not determine our research findings or the insights and recommendations of our experts. Urban scholars and experts are expected to be objective and follow the evidence wherever it may lead.

URBAN INSTITUTE · ELEVATE · THE · DEBATE

500 L'Enfant Plaza SW

Washington, DC 20024

www.urban.org