

THE FLIPPED EFFECT

Investigating the influence of
pre-class activities on exam
scores in higher education
courses using Achieve



Kara McWilliams PhD, Vice President Impact Research, Macmillan Learning

Jeff Bergin PhD, Vice President Learning Research, Macmillan Learning

Alexandra Stone, University of Connecticut

Ethics and peer review: This research complied with APA ethical standards for research. It was approved by a third-party Institutional Review Board (IRB) prior to recruitment, and then approved by individual institutional IRBs at each participating institution where required. This paper and the results herein, have been peer reviewed, revised, and approved for publication as a Macmillan Technical Report by the Impact Research Advisory Council, a panel of experts in applied research, measurement, educational technology, and the learning sciences.

Contents

6 Introduction

7 Literature review

Effectiveness of a flipped classroom

Varying delivery modes of pre-class activities

11 Achieve

12 Procedures

Research questions

Data collection

Participants and settings

Group development

18 Results

Effect of pre-class engagement on summative assessment scores in Achieve

Effect of pre-class engagement on final exam scores

Relationship between pre-class engagement and assessment performance

Relationship between pre-class engagement and perception of Achieve

Instructor and student perceptions of pre-class activities

30 Discussion

33 Limitations and future research

35 Note on data privacy

36 References



Foreword

At Macmillan Learning we are committed to providing our instructors and students with practical, actionable, and timely insights derived from studies that meet standards for educational and psychological testing. Our goal is to improve teaching and learning by enabling evidence-based decision making and to contribute to the methods and outcome research on digital learning tools. To that end, we take a comprehensive approach to measuring the effectiveness and efficacy of the digital learning tools we produce. Beginning in development, and continuing through use at scale, we partner with instructors and students to conduct studies that are appropriate for the tool's stage in the development lifecycle. Each study contributes unique and increasingly rigorous evidence to the validity and efficacy argument of that tool. Studies also produce insights into usage and engagement patterns among educational contexts that instructors might consider implementing in their own courses. This report represents one study that makes up the larger body of Achieve efficacy research. We are confident in this approach but acknowledge that measuring efficacy is complex, and we are always learning. The authors of this report, and the impact research team as a whole, welcome any comments or feedback on this report or our approach to measuring efficacy.



Acknowledgements

The work in this report reflects the efforts of many people beyond the authors. We appreciate the commitment that Macmillan Learning has made to ensuring that the tools we are developing are meeting the needs of instructors and students, and improving higher education. We thank our Chief Executive Officer, Ken Michaels for his continued support on this journey. This study would not have been a success without the partnership of Gareth Hancock and Tim Flem's Product Team, Chelsea Valentine's Technology Team, Alicia Nachman's User Experience team, and Susan Winslow's Editorial group. Our partner instructors and their students were fortunate to have the support of Marcie Smith's Customer Experience team throughout the study. And, we thank Todd Elder's design team for the production of this report. To each of the leaders at Macmillan and their talented teams, our gratitude.

In an effort to offer timely, peer reviewed insights to instructors, we are grateful to the Impact Research Advisory Council for your peer review of this report. Your guidance and critique since we began developing our approach to efficacy, ongoing insight throughout each study, and honest reviews of findings have been invaluable. Chris Dede, Michael Feldstien, Sara Finney, Suzanne Lane, Thanos Pateilis, and Elana Zieda we are indebted to you.

We also thank our colleagues on the Learning Research and Learning Analytics teams and our Chief Learning Officer, Adam Black for their reviews and helpful feedback on early drafts of this report.

Most importantly, we want to thank the instructors and students who partnered with us on this study and those who partner with us on every efficacy study we conduct. Your close partnership, keen insight, and honest feedback enable the continual optimization of the tools we develop. For the work presented in this report we particularly thank the administrators, instructors, and students at the following institutions:

Alvin Community College, Austin Community College - Cypress, Bethel University, Bunker Hill Community College, Butler County Community College California Polytechnical Institute - San Luis Obispo, Cerritos College, Delgado Community College, Drexel University, Florida State College at Jacksonville, Georgia State University - Georgia Perimeter College, Gettysburg College, Hillsborough Community College, Howard College, Northern Arizona University, Owens Community College, Pellissippi State Community College, Pomona College, Richland College, Rockhurst University, Rutgers University Busch Campus, San Francisco State University, San Jose State University, Santa Clara University, Santa Monica College, Somerset Community College, Spelman College, State University of New York College at Oneonta, University of Arizona, University of North Florida, University of Northern Texas, University of Notre Dame, University of Pittsburgh - Johnstown, University of San Francisco, University of South Carolina, Valdosta State University, and Virginia Highlands Community College.



Abstract

One strategy that has gained popularity for effectively implementing an active learning course design is a “flipped classroom” where students engage with the course material before coming to class. The literature on the effectiveness of using pre-class activities to improve learner outcomes is mixed. Consequently, we evaluated the use of pre-class activities in a new digital learning tool, Achieve, to investigate whether engaging in pre-class activities influenced assessment scores. We partnered with 40 instructors on the evaluation of pre-class activities in Achieve. Instructors chose whether or not to implement pre-class activities in their course, naturally categorizing students into “pre-class users” or “non pre-class users”. In total, 2,251 students consented to participate in the study (74% of the population), 1,372 engaged in pre-class activities and 879 did not. Groups were compared on three dependent variables: student likelihood to recommend a course using Achieve to a friend, summative assessment scores in Achieve, and in-course final exam scores. We also evaluated the effect of utilization of pre-class activities in Achieve on the dependent variables. Engagement in pre-class activities in Achieve had a significant effect on each of the dependent variables, even when prior academic performance, baseline level of motivation, and the instructor were controlled. We also found a small (based on Cohen’s classification), but significant relationship between the extent to which a student engaged in pre-class activities in Achieve and assessment scores. Finally, results uncovered evidence that instructors and students perceive pre-class activities in Achieve to contribute to positive student behaviors. And, there is evidence that use of pre-class activities significantly contributes to a student’s likelihood to recommend a course where Achieve will be used to a friend, but that it accounts for a small proportion of the variability in the outcome.



Introduction

Over the past ten years, active learning has become a popular pedagogical technique in post-secondary classrooms (Brown, Roedinger, & McDaniel, 2014). Using in-class time for higher-order cognition like application, analysis, evaluation, and creation is related to stronger learner outcomes than using the time to introduce students to content through direct instruction (Freeman, et al, 2014). One strategy for implementing an active learning course design effectively is a “flipped classroom”. In this model students engage with the course material before coming to class to gain a basic understanding of the content (Lily & Tawfik, 2015). Instructors also use this method to reduce the amount of working memory that learners have to exert, or cognitive load management, of their students by preparing them for the more demanding in-class active-learning exercises and peer-to-peer interactions (Abeysekera and Dawson 2015). Two common methods instructors use to encourage pre-class content engagement are video tutorial quizzes and reading quizzes. Video tutorial quizzes are videos describing the course content and associated formative quiz items (Basal, 2015; Graziano, 2017). Reading quizzes present students with a section of the assigned reading and associated formative quiz items. The goals of pre-class

video tutorials and reading quizzes are to offer students a basic understanding of vocabulary and core concepts as well as to support long-term learning through retrieval practice (Brown, Roedinger, & McDaniel, 2014).

Numerous studies have shown that implementing pre-class activities supports students’ academic performance, engagement, and motivation and that students and faculty have strong, positive perceptions of them (Baepler, Walker, & Driessen, 2014; Yilmaz, 2017; Zengin, 2017). However, some studies have not revealed positive impacts on student learning or student satisfaction (Kim et al., 2014; Smallhorn, 2017, Sun & Wu, 2016). In this paper, we report our findings from research investigating engagement with pre-class activities in a new digital learning tool among a variety of educational disciplines and contexts. An exploration of the effectiveness of pre-class activities—more specifically the relationship between use of pre-class activities and scores on summative assessments in the tool and final exams in the course — could extend the existing literature on pre-class activities and help educators consider how they might use this new tool in their course.



Literature review

To provide context for this study we first reviewed the literature on the use of pre-class activities in the “flipped classroom” model. Then we examined whether there was existing literature on specific types of pre-class activities used in flipped classrooms.

EFFECTIVENESS OF FLIPPED CLASSROOMS

A flipped classroom model is a method intended to increase student motivation and quality of learning by utilizing online resources such as video tutorials and gamified reading quizzes prior to attending class.

Flipped classrooms typically exhibit four components (Flipped Learning Network [FLN], 2014). First, instructors consider and communicate individual and group expectations when scheduling flexible pre-class learning. The purpose of a flipped classroom is to offer students as much time as they need to understand the material prior to engaging in in-class practical problem solving (Tawfik & Lilly, 2015). Second, instructors consider the learning cultures of their students and student subgroups within the class. Third, instructors keep a close record of individual student progression and difficulty level of the material. Lastly, instructors supply consistent feedback on each student’s performance. Once a classroom can sufficiently commit to these components of a flipped classroom, we can then examine its effect on student achievement.



Studies have demonstrated positive effects of flipped classroom methods on student achievement (e.g Sun & Wu, 2015). In a quasi-experimental study, students in a flipped classroom (experimental group) used OpenCourseWare for their pre-class educational resources, while those in the control group carried on in the conventional classroom setting. Results showed significantly higher learning achievement among the experimental group and no significant difference in teacher-student interactions. It was revealed that group work positively impacted students' academic achievements. Though greater learning achievement appears promising, data collection was based only on survey results. Students' perception of learning increased, and the flipped classroom created a richer and more dynamic physical environment for understanding class material; yet it can't be said whether flipped classrooms significantly impact students' grades from this study. In fact, one study did not find significant improvements in student achievement between flipped classrooms and traditional learning classrooms (Cabi, 2018).

Emine Cabi designed a similar study comparing the impact of flipped classrooms and traditional classrooms. Instead of surveys, exams were administered at the end of the semester to both groups, finding no significant difference in scores. While she acknowledges the benefit of students learning class material prior to class and completing assignments in class so they don't have to do them at home, she addresses two issues with the flipped classroom model as reported by students themselves. Students felt a lack of motivation to learn the material themselves and encountered issues in learning content the way it was provided

to them. Other studies have also shown a lack of beneficial outcomes to flipped classrooms (Kim et al., 2014; Smallhorn, 2017). However, there are sufficient studies supporting the benefits of the flipped classroom model against claims of ineffectiveness (Albalwi, 2018; Pierce & Fox, 2012; Talley & Scherer, 2013; Zengin, 2017).

Studies that have shown positive impacts of flipped classroom models are largely found within mathematics classrooms (Freudenthal, 1983). Contrary to previous studies mentioned, one study compared the outcomes of students in a "flipped" classroom and those in a traditional classroom and found significant differences in student achievement of those in a flipped calculus classroom (Albalawi, 2018). Another study found significantly higher Algebra I grades among students enrolled in the flipped classroom versus the traditional learning environment (Wiginton, 2013). While these same students enjoyed learning at their own pace, some struggled with falling behind. The flipped classroom model is not a cookie-cutter learning environment with the same implementation in all classrooms. Its success is dependent on many variables including student motivation and the quality of pre-class activities. Among the most common pre-class activities are video tutorials, administered through the institutions LMS or third-party educational content creators.



VARYING DELIVERY MODES OF PRE-CLASS ACTIVITIES

VIDEO TUTORIALS. The theory that video tutorials should replace post-lecture assignments has grown popular for the benefit of affording more class time to differentiated education and practical problem solving (Davies et al., 2013). Additionally, pre-class video tutorials allow students to learn at their own pace, with the option to rewatch or reiterate the material before engaging in active group learning in class. It has also shown to be beneficial for students to have access to past video tutorials, covering class material that has been covered in the past, allowing students to double check their knowledge of concepts they previously found difficult (Lily & Tawfik, 2015). One study utilized “vodcasts”, or video podcasts, as a substitute for in-class lectures (Pierce & Fox, 2012). The vodcasts were videos of professors lecturing, which could be accessed through iTunes. After watching these pre-class vodcasts, students

engaged in group discussion and problem solving, which showed significant improvements in academic achievement.

Although these studies suggest positive outcomes from pre-class tutorial video in place of in-class lectures, some studies show a contrast in student preference for video tutorial learning (Bishop & Verleger, 2013). Students reported a preference of in-class lectures over video lectures, yet they preferred interactive in-class activities. The challenge is achieving the best of both worlds while optimizing class time. While students must assume the responsibility of their education by actively watching and taking notes on these video tutorials, professors must carefully plan appropriate videos and presentations outside of class, as it is not a substitute for in-class lectures but a supplement to in-class activities (Zengin, 2017).



ADAPTIVE, GAMIFIED READING QUIZZES.

Adaptive gamified reading quizzes ask students to read a portion of a text and answer items that are developed to adapt to student correct or incorrect responses. The “gamified” nature is introduced when students have some target or goal that they have to hit to receive credit. Gamified quizzing is an additional and equally successful pre-class assessment to intrinsically motivate students’ learning processes outside of class. In a five-year longitudinal study, researchers compared the first three years of non-gamified assignments completed outside of class and two successive years of experimental gamified quizzing conducted outside of class (Barata et al., 2013). Results showed significantly higher grades among students participating in the gamified assignments outside of class. Additionally, those students showed improvements in attention to reference materials, online participation, and proactivity (Barata et al., 2013).

Not only do gamified activities improve students’ retention of course content, it has also shown to improve student work ethic and self-reliance (Latulipe et al., 2015). This finding could remedy the issue of students’ lack of motivation in learning course content prior to in-class activities where they may fall behind. In a mixed-method pilot study showing significantly higher scores among students in a flipped-classroom, researchers highlighted the benefits of friendly competition and students’ social engagement through gamified activities completed outside of class (Zainuddin, 2018). The same study collected student survey responses, in which they reported positively perceived competence, autonomy, and relatedness when completing gamified activities outside of class. Another study concluded that student achievement is optimized when (1) the

instructor is viewed as a mentor and (2) there is a balance of extrinsic and intrinsic motivation within a learning environment, both in-class and outside of class (Dicheva & Dichev, 2016). A combination of the flipped classroom model and gamified quizzes/activities may be the most fitting combination, placing higher expectations on students to commit sufficient hours of learning outside of class while offering creative methods of course review through potentially competitive games in which students motivate each other to perform better and learn more.

SUMMARY

Effective flipped classrooms allow students to learn at their own pace, contribute more effectively during in-class activities, and boost social engagement between students during group-work activities. There are multiple factors that can predict the success of a flipped classroom, such as sub-groups of students engaged in the course topic, previous online class experience, and academic rigor regularly expected by their institution. Future research should begin to target more specific student level variables, examining why some case studies show significant or insignificant improvements in students’ grades due to the flipped classroom model. With this information, instructors can better determine if a flipped classroom model will benefit students in their course.

With a solid grounding in the literature, the next section describes Achieve, the digital learning tool that this study investigated and the specific pre-class activities examined within Achieve.



Achieve

Achieve is a digital learning solution developed for higher education courses. It provides a connected suite of course tools designed to give instructors choice, with flexible recommendations for optimal learning paths based on the learning sciences. The key principles that Achieve is built on include: everyone has the potential to learn, each learner starts at a different place and learns at their own pace, cognition can be enhanced through technology, an instructor's pedagogy matters, learning is a social activity, and students should be empowered to manage their learning. Achieve is built on learning science foundations developed based on research from experts on how students learn and how instructors and students can achieve desired outcomes. Based on these foundations a learning model was constructed to act as a blueprint for the design of Achieve; as part of the learning model pre-class activities were provided as one of the resources that instructors might choose to implement.



Procedures

This research complied with American Psychological Association ethical standards for research. It was approved by a third-party Institutional Review Board (IRB) prior to participant recruitment, and then approved by instructor participant's individual institutional IRBs where required.

In the Spring 2019 semester 40 instructors across five disciplines (Biology, Calculus, Chemistry, Composition, and Economics) agreed to participate in an evaluation of Achieve, a new digital learning platform, before it was being used at scale. Instructors and students received Achieve free of charge to use. All students were required to use Achieve in their course because it was the curricular material their instructor selected, but they were not required to participate in this study. Interested students were required to actively consent to participate. The 2,251 students in the evaluation study made up 74% of all students enrolled in participating courses.

Prior to the beginning of the semester Instructors were required to complete a thirty minute training on Achieve. During training instructors were offered suggestions for best-practice implementation based on learning science research – including assigning pre-class activities - but implementation patterns were not mandated as part of the evaluation. The only implementation requirement was that Achieve had to be the primary curricular material used that semester. Pre-class activities in Achieve were available to assign across all five disciplines represented in this study. Two student groups naturally emerged from self selected implementation patterns (1) instructors who did not implement pre-class activities in their course (2) instructors who elected to assign either pre-class video tutorials, pre-class adaptive quizzing, or both pre-class video tutorials and adaptive quizzes in their course.



PRE-CLASS VIDEOS. Students are asked to view brief, five to seven-minute videos. In most cases there was an associated set of non-adaptive quiz items related to the content in the video. After playing the video once, students were permitted to view the content in subsequent replays without any penalty. Where quiz questions were provided, students were given multiple attempts to complete the questions correctly and could refer back to the video while responding to the questions. The total points that a student received for each question decreased with each incorrect attempt. Students were also given the option to receive hints about the answer to the question which decreased the total points available to them per item. In all courses represented in the study student performance on the quizzes did factor into their overall course grade. The proportion of the final course grade that pre-class videos accounted for ranged from 2% to 5%.

GAMIFIED, ADAPTIVE READING QUIZ. Instructors identified a “target score” for the reading quiz that students were expected to reach, and students had to continue answering questions until they reached that score. If a student answered a question correctly on the first try, they earned the total number of points for that item; each subsequent attempt decreased the total number of points earned for that item. If a student got a quiz item wrong they were presented another item with similar content and similar level of difficulty; if they answered the item correctly they were presented an item assessing different content with slightly more difficulty. Students were also given the option to receive hints about the answer to the question which decreased the total points available to them per item. Students could also request to be taken directly to the portion of

the etext where the content needed to correctly answer the question could be found. Student scores did not decrease if they requested to “refer to the text”. Students could also select the “show answer” tab, but earned no points for the item if they selected this option. The adaptive reading quiz was graded for completion. Students received full credit (100) if they reached target score and no credit if they didn’t (0). Students were permitted to continue answering questions beyond the target score but it did not impact their overall score. Performance on the adaptive quizzes was factored into a student’s overall grade in all courses in the study. The proportion of the final course grade that the reading quizzes accounted for ranged from 2% to 10%.

Achieve was built on the premise that having students engage with content before coming to class enables instructors to replace direct instruction with active learning and higher order thinking. And, that video and reading quizzing promotes retrieval practice. Based on this development the hypothesis was that engagement with pre-class activities would support understanding, information retention, and retrieval. Therefore in this study the authors sought to validate that student engagement with pre-class activities in Achieve would support future academic performance both within Achieve (because the content would be directly aligned) and outside of the platform in the course (because of the importance of measuring overall learning).

We expected that if we validated the relationship between pre-class activity engagement and assessment performance, findings could support the implementation decisions instructors make when using Achieve. Findings could also inform training, professional development, and possibly product improvements that help promote implementation of pre-class activities. If the research suggested that engagement in pre-class activities did not support subsequent academic performance, the findings would support a re-evaluation of the pre-class activities, the quality of the content and items, and the alignment between the content in the pre-class activities and summative assessments.



RESEARCH QUESTIONS

Based on how instructors decided to implement Achieve, students fell into one of two groups (1) were assigned pre-class activities for course credit - “pre-class” (2) were not assigned pre-class activities - “no pre-class”.

This study addressed five research questions designed to help educators better understand whether instructors and students value pre-class activities in Achieve, and whether implementing pre-class activities in Achieve positively influences student learning.

- 1.) *What is the effect of engaging in Achieve pre-class assessments on summative assessment scores in Achieve when prior academic performance, baseline level of motivation, and instructor are controlled?*
- 2.) *What is the effect of engaging in Achieve pre-class assessments on final exam scores when prior academic performance, baseline level of motivation, and instructor are controlled?*
- 3.) *Is more utilization of pre-class activities in Achieve related to performance on summative assessments in Achieve and final exam scores?*
- 4.) *Does use of pre-class activities in Achieve influence student perceptions of Achieve?*
- 5.) *What are student and instructor perceptions of pre-class activities in Achieve?*

DATA COLLECTION

Data were collected for a mixed-methods analysis. Student and instructor surveys were administered at the beginning and end of the semester, instructors completed weekly implementation logs, and instructor interviews were conducted mid-semester. Product usage data were extracted from the Achieve platform on a weekly basis and at the end of the study, and student records were shared by instructors at the end of the semester. Data were matched across sources, and descriptive and empirical analyses were conducted. A complete description of the collected data follows.

STUDENT PRE-SURVEY. A link to an online survey that asked students to first consent to participate in the study and then report their background and demographic characteristics was shared by their instructor during the first two weeks of the course. The survey captured data on student comfort with technology, student sentiment toward technology use in the classroom, value of digital tools in the classroom, academic behaviors outside of class, classroom behavior, and sentiment toward the course. Students were also asked to report their major, whether they were taking this course as part of their major requirement, high school grade point average, whether they took the SAT and/or ACT, and their scores on each section, as well as various demographic data. These data were collected as potential moderators of the relationship between use of pre-class activities and academic performance and were used in the analyses of the research questions.¹

INSTRUCTOR PRE-SURVEY. A survey was administered to instructors online during the first month of the courses. The survey included a scale that measured acceptance of technology and included items about comfort with technology, perceptions of technology in the classroom, intended implementation of Achieve, intended implementation of other publisher-provided digital learning tools or open educational resources, previous experience with Achieve, and general early perceptions of Achieve. These data were used to control for instructor characteristics and to better understand intended implementation of Achieve.

1: Although students self-report their measures of prior academic performance, we can have confidence in the reliability of the scores based on previous research. [Shaw and Matter \(2009\) examined the reliability of self-reported HSGPA and found a correlation of 0.74 and in a 2005 meta-analysis, Kuncel, Credé, and Thomas found a correlation of .82 between actual and self-reported SAT scores.](#)



INSTRUCTOR WEEKLY IMPLEMENTATION LOGS.

An online survey was sent to instructors at the end of each week. The survey asked instructors to report how they implemented Achieve in the previous week (which features and components they used), how much time various activities took them, their perception of Achieve that week, any benefits or challenges of using Achieve, and any other information that would help us understand usage that week (e.g. whether class was canceled for inclement weather). These data were used to track ongoing actual implementation and how that was related to perception.

INSTRUCTOR INTERVIEWS. An instructor interview protocol was developed that gathered information on how an instructor was implementing Achieve, why they decided to implement it in that way, their perceptions of Achieve, and their perceptions of how their students were accepting Achieve. Probes were developed based on the responses provided in implementation logs and in real-time based on responses to questions in the interview protocol. These data were used to better understand why an instructor chose to assign pre-class activities or not assign pre-class activities.

INSTRUCTOR POST-SURVEY. A survey was administered online during the last two weeks of the course to instructors. The survey included a scale that measured instructor ability to assess student understanding; a scale to assess active learning in the classroom; a scale to measure classroom challenges; a scale to measure student behavior, their implementation of Achieve and their perceptions of Achieve; a System Usability Scale²; the Net Promoter Score, and a likelihood of adoption scale. These data were used to measure whether there were systematic differences between instructors who assigned pre-class activities and those who didn't.

STUDENT POST-SURVEY. A survey was administered online during the last two weeks of the semester. The survey asked students to share demographic data, personal device data, how they used Achieve, their perceptions of Achieve, their engagement in the course³, their satisfaction with the course, a System Usability Scale, and a Net Promoter Score. These data were used to measure whether there were systematic differences between students who engaged in pre-class activities and those who didn't.

PRODUCT USAGE DATA. The following data were extracted from the Achieve platform for consenting students: student name, student email, each activity that an instructor assigned, assignment date and due date, whether student accessed each activity, student progress on each assigned activity, student completion of each assigned activity, student performance on each assigned activity, student access of unassigned activities, student progress on unassigned activities, student completion of each unassigned activity, student performance on each unassigned activity. These data were used to measure actual instructor implementation and student usage.

STUDENT RECORDS. Instructors were asked to share the following course performance data for consenting students: homework scores, quiz scores, exam scores, final exam scores, final course grades and percentages, attendance rate, and participation scores. Instructors were not asked to change their regular course performance methods, so some data were not available for all students. For example, some instructors did not score homework or give quizzes, so they only reported exam scores and final course grades. And, not all class records were provided in the same metric, so only grades that could be reliably compared were included in the analysis. In this report, final exam scores were the only student record used.

2: Brooke, J. (1986). System Usability Scale. Digital Equipment Corporation.

3: Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *The Journal of Educational Research*, 98 (3), 184-191.



PARTICIPANTS AND SETTINGS

To investigate the relationship between engaging in pre-class activities on student outcomes, we investigated data collected from 2,251 undergraduate students enrolled at 38 institutions in 42 courses among 40 instructors.

There was variability in institution and course type in the sample. Across institutions 42% were two-year and 58% were four-year, the majority of institutions (76%) had more than 2,000 undergraduate students enrolled, and based on the Carnegie Classification of Institutions of Higher Education, 45% were more selective, 34% were moderately selective, and 16% were less selective. Among courses the largest proportion (21%) were Calculus I courses, 19% were Chemistry I courses, 17% were Non-majors Biology courses, 17% were Microeconomics courses, 14% were Macroeconomics courses, 12% were Composition courses.

Instructor background and demographic information was self-reported on the baseline and end of semester surveys and included years teaching (less than five years = 45%, six to fifteen years = 15%, more than 15 year = 40%), comfort level using technology (not comfortable at all = 0%, somewhat comfortable = 22%, comfortable = 74%, very comfortable = 4%).

Student demographic information was self-reported on the baseline and end of semester student surveys and is presented in Table 1.

GROUP DEVELOPMENT

Platform usage data were extracted and analyzed to develop the groups that would be used in this study. An instructor was considered a “pre-class assigner” if they assigned at least one pre-class activity in more than 10% of the weeks in the semester. If an instructor never assigned a pre-class activity or assigned a pre-class activity in 10% or fewer of the weeks in the semester they were considered a “non-assigner”. The decision was made to code instructors who assigned a pre-class activity in less than 10% of weeks as a “non-assigner” because the average proportion of weeks that an “assigner” assigned a pre-class activity was 64.7%, so a student would not have had the same exposure if they were assigned it less than 10% of the semester weeks. Note, that an instructor may have assigned one of these activities as a “post-class” activity. In those cases the assignment of the activity did not contribute to a pre-class assignment. In total, 26 instructors were “pre-class assigners” and 14 instructors were “non-assigners”. The number of pre-class activities assigned in a week by assigning instructors ranged from (0-4). Pre-class activities were either short videos with associated quiz items or adaptive reading quizzes. Both modes of pre-class activities were completed online before coming to the class in which that content would be covered.

Students were then classified as either a “pre-class user” or “non pre-class user” based on which course they were enrolled in (i.e. which instructor they had). In total, 1,372 students engaged in pre-class activities and 879 did not. Table 1 compares the descriptive statistics betw the two groups of students.



Table 1. Student group and descriptive comparison

	All	Pre-class user (n=1,372)	Non pre-class user (n=879)
Discipline			
Biology	9.56	13.34	3.35
Calculus	10.42	8.82	13.04
Chemistry	35.73	24.42	54.31
Economics	14.61	51.60	6.70
Composition	9.69	1.82	22.61
Year in college			
Dual enrolled	0.83	0.30	1.71
First	57.79	63.05	49.21
Second	20.55	18.77	23.44
Third	10.53	8.15	14.41
Fourth	3.76	3.96	3.42
Fifth	0.83	0.60	1.22
Other	5.69	5.16	6.59
Eligible for federal financial aid			
Eligible	64.23	62.34	68.91
Ineligible	35.77	37.66	31.09
First generation			
Yes	22.77	21.03	25.48
No	77.23	78.97	74.52
Gender			
Male	44.56	43.41	47.41
Female	54.84	56.38	51.04
Prefer not to say	0.60	0.21	1.55
Taking the course as discipline requirement			
Yes	72.01	71.03	74.22
No	27.99	28.87	25.78
Traditionally underrepresented	27.94	26.20	31.50
HSGPA*	3.65	3.73	3.51
Baseline level of motivation to succeed	0.75	0.75	0.77
Average summative assessment score*	75.75	77.86	69.20
Average final exam score*	64.13	66.15	59.36

*denotes statistically significant difference in average at the p = 0.05 level



Results

RESEARCH QUESTION 1

What is the effect of engaging in pre-class activities in Achieve on average summative assessment score in Achieve when prior academic performance, baseline level of motivation, and instructor are controlled?

We hypothesized that being assigned and engaging in pre-class activities in Achieve would positively influence summative assessment scores in Achieve. We expected that all students would benefit from class preparedness because we hypothesized: that class preparedness would help close skills gaps while engaging and motivating students, that students coming to class with an introduction to the content would reduce the need for direct instruction, and that it would enable higher order thinking and more active learning during course time. We also hypothesized that pre-class activities would influence a testing effect.

The dependent variable was average summative assessment score in Achieve. Summative assessment scores in Achieve are aligned to the pre-class material because they assess the same content, and the delivery and scoring are standardized across disciplines.

Data were reduced to students who had valid data on all variables in the models (n=1,711).



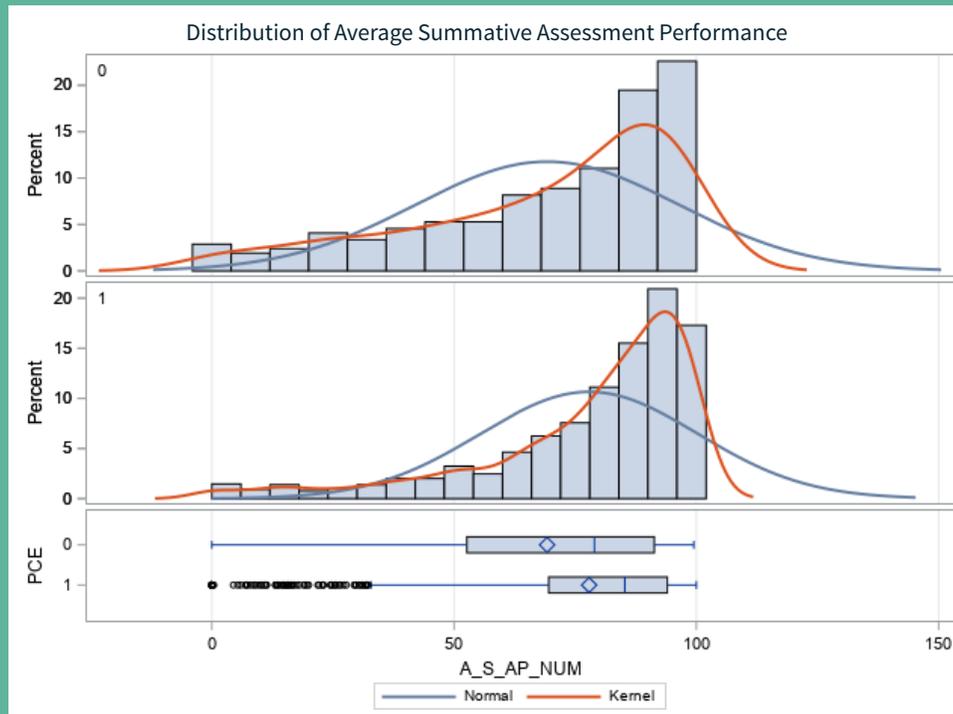
First, we descriptively examined the difference in average summative assessment scores in Achieve by students who had been assigned and engaged in pre-class assessments (n=1,295; mean percentage points = 77.86, SD = 22.46) and those who had not been assigned pre-class activities (n= 416; mean percentage points = 69.20, SD = 27.10). Pre-class assessment students earned 8.7 percentage points higher, on average. When group differences were compared statistically, there was a statistically significant difference. Figure 1 presents the distribution of average summative assessment scores in Achieve by group.

A significant correlation between the dependent variable, average summative assessment score in Achieve and high school grade point average (HSGPA) and baseline level of motivation were found .28 ($p < .001$) and .08 ($p = .0028$), respectively, and students were grouped within instructors who likely had systematic differences in both characteristics and pedagogical approaches.

Consequently, we implemented a hierarchical linear model and controlled for HSGPA and baseline level of motivation to investigate the net relationship between use of Achieve pre-class assessments and summative assessment scores in Achieve when controlling for related covariates.

A hierarchical linear model was calculated using PROC MIXED in SAS to examine the interaction effects in this model. The first model investigated was the unconditional model with no predictors to assess between instructor variation in summative assessment performance in Achieve. The intraclass correlation coefficient was calculated as $ICC = (163.27)/(163.27+474.87) = 0.260$. This finding suggests that about 26% of the variability in Achieve summative assessment score was attributed to the instructor that the student had, and 74% of the variability was attributable to the student. This finding aligned with our hypothesis and confirmed that we should employ a hierarchical linear model.

Figure 1. Distribution and box plot of average summative assessment scores in Achieve by group





We continued the model building process by including student prior academic performance and baseline level of motivation to succeed in the course as fixed effects, then including them as random effects. Results of the second model demonstrated that the model was significant and both HSGPA ($p < 0.0001$) and baseline level of motivation were significant ($p = 0.0003$). We then added the variables as random effects to determine if the influence on summative assessment scores in Achieve varied among instructor. The model remained significant as did HSGPA ($p < 0.0001$) and baseline level of motivation ($p = 0.0270$).

We continued the model building process by adding instructor's experience teaching and whether the student was assigned and engaged in pre-class activities in Achieve. An instructor's experience using technology did not significantly contribute to model fit so it was removed from the model and the model was rerun. Results from this model and summary results of this model building process are presented in Table 2.

We evaluated the change in AIC and BIC and concluded that Model four was the best fitting model. And, given that the inclusion of whether a student was assigned and engaged in pre-class activities in Achieve emerged as significant we have evidence to conclude that the use of pre-class activities in Achieve influenced average summative assessment scores in Achieve. More specifically, when controlling for student HSGPA and baseline level of motivation the average summative assessment scores in Achieve for students who utilized pre-class assessments was significantly different from the average summative assessment for students in courses where they were not. Given these findings, we can conclude that assigning pre-class activities in Achieve significantly supports better performance on summative assessments within Achieve.

Table 2. Estimates for two-level organizational linear model predicting summative assessment scores in Achieve

	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Intercept	72.403* (2.36)	30.03* (4.87)	30.86* (4.74)	6.84 (5.74)
HSGPA		10.62* (1.14)	10.38* (1.26)	10.38* (1.13)
Motivation		4.89* (1.34)	4.99* (1.78)	5.01* (1.77)
Pre-class				34.89* (3.93)
Error variance				
Level-1	474.87* (16.40)	426.06* (16.13)	423.17* (16.22)	397.45* (15.44)
Level-2 intercept	163.27* (46.43)	145.75* (42.86)	57.62 (74.96)	235.06* (100.57)
HSGPA			5.05 (5.02)	0.311 (3.61)
Motivation			19.69 (30.8)	19.7511 (3.07)
Model fit				
AIC	15488.4	12802	12804.1	12734.9
BIC	15493.6	12810.5	12816.1	12748.6

Note: * denotes statistical significances, $p < .05$; ICC = .25, error variances reported in Table 1
 Values based on SAS PROC MIXED. Entires show parameter estimates with standard errors in parentheses
 Estimation Method = ML; Satterthwaite degrees of freedom



RESEARCH QUESTION 2

What is the effect of taking pre-class assessments on final exam scores when prior academic performance, baseline level of motivation, and instructor are controlled?

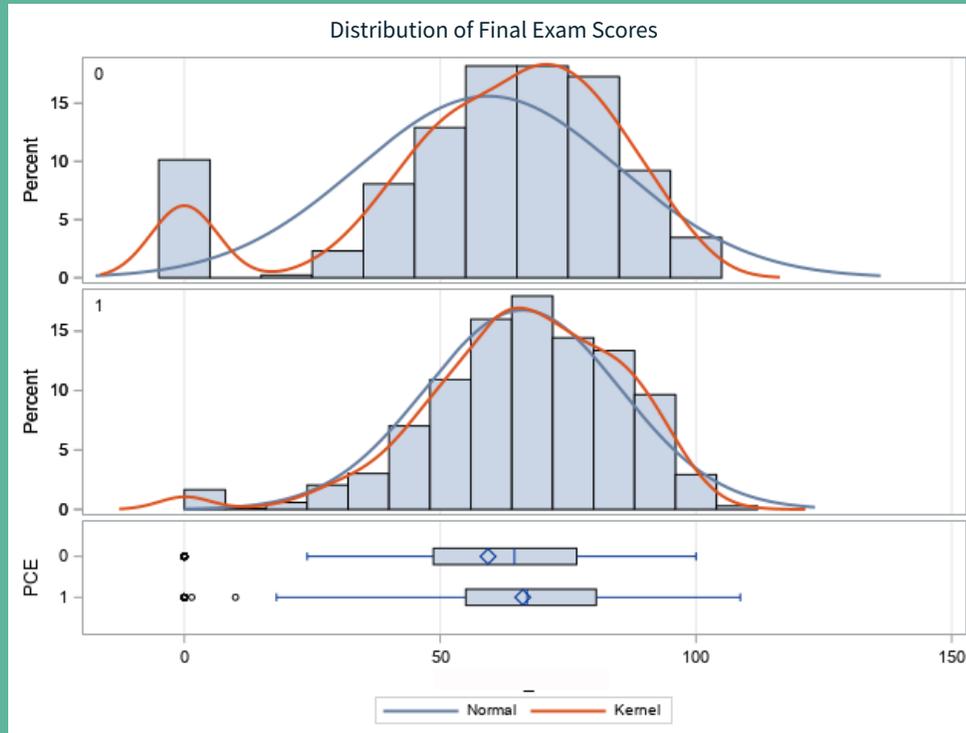
We hypothesized that being assigned and engaging in pre-class activities in Achieve would positively influence in-course final exam scores. We expected that all students would benefit from class preparedness because we hypothesized: that class preparedness would help close skills gaps while engaging and motivating students, that students coming to class with an introduction to the content would reduce the need for direct instruction, and that it would enable higher order thinking and more active learning during course time. We also hypothesized that pre-class activities would influence a testing effect.

First, we descriptively examined the difference in final exam scores by students who had been assigned and engaged in pre-class assessments (Mean = 66.15, SD = 19.04) and those who had not been assigned pre-class activities (Mean = 59.36, SD = 25.57). Pre-class assessment students earned 6.78 percentage points higher, on average. When compared statistically, the difference was significant. Note that there was no mandated standardization of final exam scores, which was a

limitation of the study and why we also examined summative assessment scores in research question 1. Figure 2 presents the distribution of final exam scores by group. Note, the proportion of final exam grades where students earned less than 1 percentage points. We validated with instructors that those grades were accurate and therefore kept them in the analytic sample.

The correlation between the dependent variable, final exam score and high school grade point average (HSGPA) was significant .22 ($p < .0001$) so we controlled for HSGPA. The correlation between final exam score and student baseline level of motivation was not significant $R^2 = 0.016$, $p = 0.5630$ but we theorized, based on previous research that baseline level of motivation was related to the score a student earned on their final exam and thus included it as a covariate in the model. In addition, students were grouped within instructors who likely had systematic differences in both characteristics and pedagogical approaches so we employed a hierarchical linear model.

Figure 2. Distribution of final exam scores by group



A hierarchical linear model was calculated using PROC MIXED in SAS. The first model investigated was the unconditional model with no predictors to assess between instructor variation in final exam performance. The intraclass correlation coefficient was calculated as $ICC = (116.89)/(116.89+369.51) = 0.240$. This finding suggests that about 24% of the variability in final exam score was attributed to the instructor that the student had, and 76% of the variability was attributable to the student. This finding aligned with our hypothesis and confirmed that we should employ a hierarchical linear model.

We continued the model building process by first including student prior academic performance and baseline level of motivation to succeed in the course as fixed effects, then including them as random effects. Results of the second model demonstrated that the model was

significant. When considered individually HSGPA was significant ($p < 0.001$) but baseline level of motivation was not significant ($p = 0.3775$). We then added the variables as random effects to determine if the influence on summative assessment scores in Achieve varied among instructor. The model remained significant as did HSGPA ($p < 0.001$) and again level of motivation was not significant ($p = 0.3780$).

We continued the model building process by adding instructor's experience teaching and whether the student was assigned and engaged in pre-class activities. An instructor's experience using technology did not significantly contribute to model fit, so it was removed from the model and the model was rerun. Results from this model and summary results of this model building process are presented in Table 3.



We evaluated the change in AIC and BIC and concluded that Model four was the best fitting model. And, given that the inclusion of whether a student was assigned and engaged in Achieve pre-class assessment emerged as significant we have evidence to conclude that the treatment of pre-class activities in Achieve influenced average final exam scores. More specifically, when

controlling for students HSGPA, the average final exam scores for students in courses where Achieve pre-class assessments were assigned were significantly different from the average final exam scores for students in courses where they were not. Given this evidence, we can conclude that assigning pre-class activities in Achieve supports better performance on final exams in their course.

Table 3. Estimates for two-level organizational linear model predicting final exam score

		Model 1	Model2	Model 3	Model 4
	Fixed effects				
Intercept		72.403* (2.36)	35.33* (4.85)	35.33* (4.84)	23.48* (5.62)
HSGPA			8.14* (1.14)	8.14* (1.14)	8.03* (1.13)
Motivation			1.16 (1.32)	1.16 (1.32)	1.09 (1.30)
Pre-class					19.47* (3.61)
	Error variance				
Level-1		369.51* (13.82)	352.98* (14.35)	352.98* (14.35)	341.24* (13.95)
Level-2intercept		116.89* (35.54)	120.47* (37.04)	120.52* (37.06)	229.77* (79.62)
HSGPA				0	0
Motivation				0	0
	Model fit				
AIC		12839	10850.7	10850.7	10828.2
BIC		12844.2	10859.2	10859.2	10838.5

Note: * denotes statistical significances, $p < .05$; ICC= .25, error variances reported in Table 2
 Values based on SAS PROC MIXED. Entires show parameter estimates with standard errors in parentheses
 Estimation Method = ML; Satterthwaite degrees of freedom



RESEARCH QUESTION 3

Is higher engagement in assigned pre-class activities related to performance on assessments?

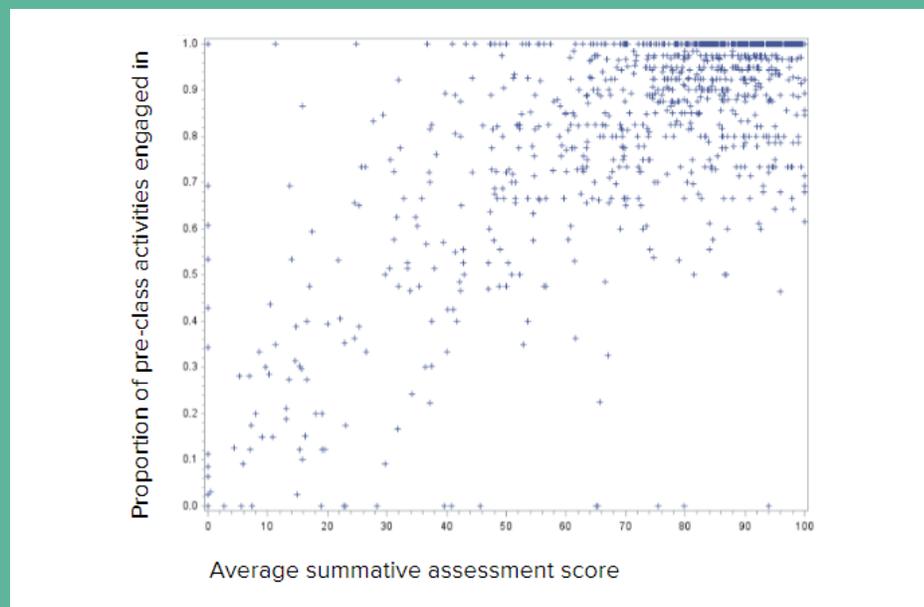
We hypothesized that there would be a positive correlation between the proportion of assigned pre-class activities in Achieve that a student engaged in and their average summative assessment scores in Achieve for the following reasons: The more pre-class activities in Achieve that a student engages with the more exposure to material the student has and, the more pre-class activities in Achieve that a student engages in the more they will be able to engage in higher-order cognition work in the classroom rather than direct instruction because they will have received introductory information in the pre-class activity.

Students in the “pre-class user” group (n=1,372) were used to answer this research question. We first investigated performance in Achieve. The two

variables that were examined in this analysis were “percent of assigned pre-class activities a student engaged in” and “average summative assessment score”. The sample was reduced to students with valid metrics on both variables (n=1,316).

Results of the Pearson correlation indicated that there was a strong (based on Cohen (1988) classification), statistically significant positive correlation between the proportion of pre-class activities in Achieve a student engaged in and their average Achieve summative assessment score ($r(1,316) = .71, p < .0001$). This finding suggests that the more assigned pre-class activities in Achieve that a student engages in the higher their average Achieve summative assessment score.

Figure 3. Correlation between the percent of assigned pre-class activities in Achieve that a student engaged in and their average Achieve summative assessment score



Note: n=1,316



To further explore this relationship descriptively, we disaggregated proportion of engagement with Achieve pre-class assessments into bands and examined the average Achieve summative assessment score by engagement band.

We then investigated the relationship between engagement with pre-class activities in Achieve and final exam scores. Students in the “pre-class user” group (n=1,372) were used to answer this research question. The two variables that were examined in this analysis were “percent of assigned pre-class activities a student engaged in” and “final exam score”. The sample was reduced to students with valid metrics on both variables (n=1,040).

Results of the Pearson correlation indicated that there was a weak (based on Cohen’s 1988 classification), but statistically significantly positive relationship between the proportion of pre-class activities in Achieve a student engaged

in and their final exam score ($r(1,040) = .33, p < .0001$), as seen in Figure 5. The correlation with final exam scores being weaker than summative assessment scores in Achieve is not surprising given that the content assessed on final exams was not as aligned as that which was assessed on summative assessments in Achieve. The weak correlation notwithstanding, the significance suggests that the more pre-class activities assigned in Achieve that a student engages in, the higher their final exam score.

The results of the incremental analysis are presented below. The same relationship can be seen when examining final exam score as observed when investigating average summative assessment score, however it is important to note the small sample size when disaggregated by engagement band, so findings should be interpreted with caution and as suggestive.

Figure 4. Average summative assessment performance in Achieve by assigned Achieve pre-class activity engagement band

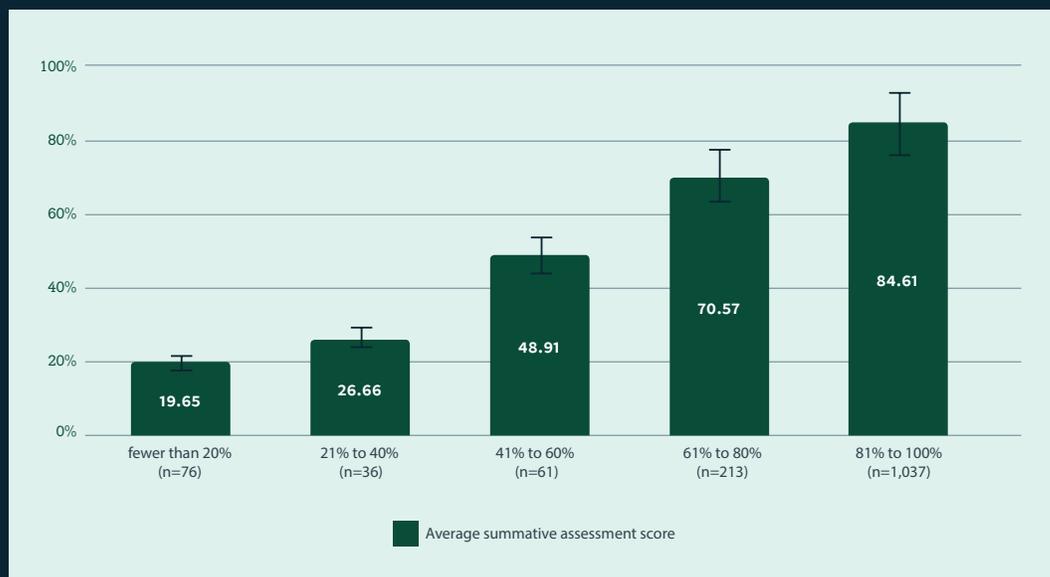




Figure 5. Correlation between the percent of pre-class activities in Achieve that a student engaged in and their final exam score

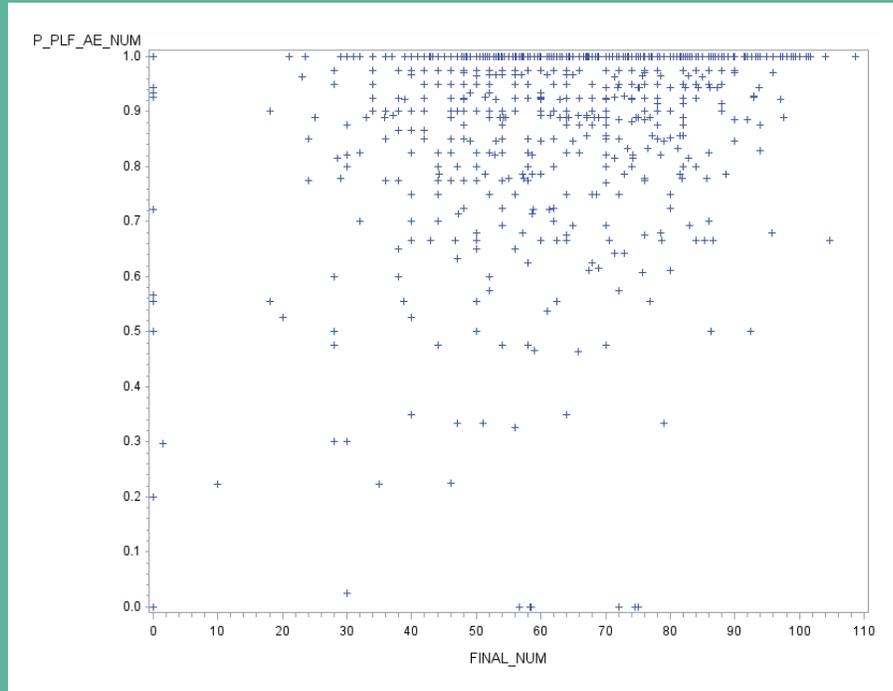
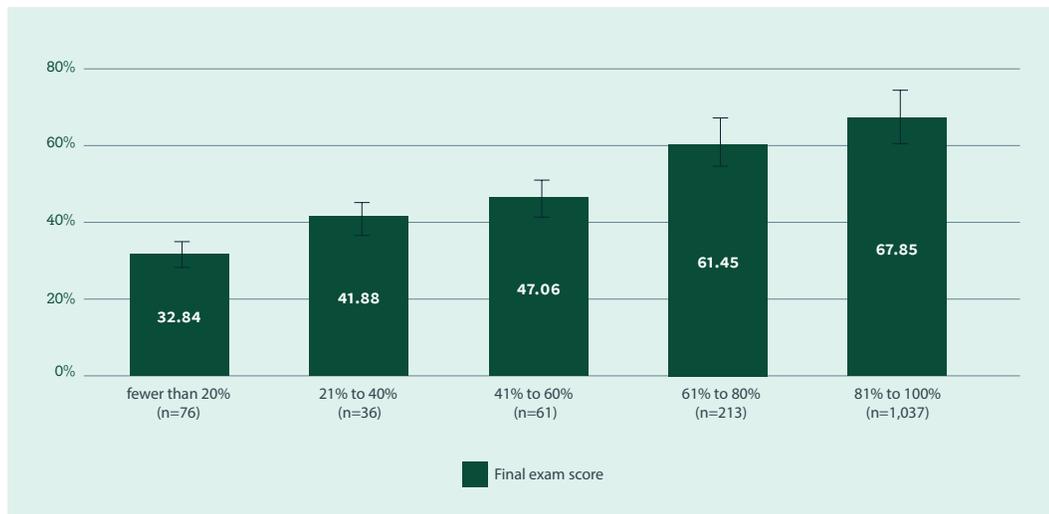


Figure 6. Average final exam score by engagement with assigned pre-class assessment band





RESEARCH QUESTION 4

Does use of pre-class activities in Achieve influence student perceptions of Achieve?

We hypothesized that students who engaged in pre-class activities would have a more positive perception of Achieve than students who didn't use pre-class activities because: they would enjoy the pre-class activities, they would come to class more prepared, giving them more confidence to actively participate in class and, they would realize higher course grades as a result of use of Achieve.

First, one item on the post-survey asked students to rate on a scale of 0-10 whether they would recommend the course to a friend (NPS), which was used as the outcome variable studied. A multiple linear regression was calculated to predict NPS based on whether or not they engaged in any pre-class activities, and to control for variables related to satisfaction we included HSGPA, baseline level of motivation to succeed in the course, whether the student was taking the course as a major requirement, and the student's level of comfort using technology in the classroom. A significant regression equation was found ($F(1,022) = 6.65, p < .0001$) with an R^2 of 0.032 but none of the covariates were significant predictors of NPS score.

The results suggest that whether students engaged in pre-class activities significantly contributed to their likelihood to recommend a course to a friend if they knew Achieve was going to be used, however it is important to note that the $R^2 = 0.0320$, suggesting that only 3% of the variance in student perception could be attributed to using pre-class in the course. So while the finding is statistically significant, there is likely very little practical significance and the reader should be cautious of the interpretation as it suggests many other factors are related to whether a student would recommend a course using Achieve.

To further investigate this research question we compared other survey responses between pre-class users ($n=1,373$) and non-users ($n=879$). We investigated whether there was a statistically significant difference in student perception of ease of use to determine whether adding this component would make the program more complex for students. Students were asked to respond to a single ease of use question rated on a scale of 1 = "very difficult" to 4 = "very easy". Results show that pre-class users rated Achieve significantly easier to use than students who did not use pre-class activities. Students also



completed a system usability scale, and again pre-class users rated the system significantly easier to use than non pre-class users. These findings suggest that the addition of pre-class activities into a curriculum did not complicate use of the program.

Students were also asked to rate their level of agreement on four items on a scale of 1 = “strongly disagree” through 4 = “strongly agree”.

Students who were pre-class users rated “Achieve helped me gain mastery of the course material” significantly higher than no pre-class users. There was no significant difference in ratings of “Achieve helped me comprehend the course material”, “Achieve helped me fill skills gaps that I had” or “Achieve helped me recall concepts”. All results can be found in Table 4.

Table 4. Perceptions of Achieve by whether pre-class activities were assigned

	All	Pre-class user	Non pre-class user
Ease of use*	3.05	3.09	2.96
System usability scale*	63.71	64.98	60.54
Achieve helped me gain mastery of the course material*	2.87	2.90	2.79
Achieve helped me comprehend the course material	3.05	3.04	3.09
Achieve helped me fill skills gaps that I had	2.35	2.29	2.52
Achieve helped me recall concepts	3.00	3.00	3.02



RESEARCH QUESTION 5

What are instructor and student perceptions of pre-class activities in Achieve?

We hypothesized that instructors and students would have strong, positive perceptions of pre-class activities. We hypothesized that instructors would perceive pre-class activities to help students stay on track with the reading, come to class prepared to participate, enable more active learning in the classroom, and support stronger academic performance in the course. We hypothesized that students would enjoy the pre-class activities, come to class more prepared, giving them more confidence to actively participate in class and realize higher course grades as a result of use of Achieve.

On the post-survey instructors and students were asked to report their perception of pre-class activities. Instructors who assigned pre-class activities and students of instructors who assigned pre-class activities were asked the same set of questions about their perception of pre-class activities using slightly different wording. Instructors were asked to rate on a scale of 1 = “strongly disagree” to 4 = “strongly agree” the extent to which they agreed that pre-class activities in Achieve helped their students stay on

track with the reading this semester, the average rating was 3.00 among 26 instructors. Students were asked to rate on the same scale the extent to which they agreed that pre-class activities in Achieve helped them stay on track with their reading this semester. The average rating among 964 students was 2.96.

Instructors and students were also asked the extent to which they agree that pre-class activities in Achieve gave students a basic understanding of concepts and both instructors and students rated tended to agree (3.03 and 3.08, respectively). When asked whether pre-class activities provided a basic understanding of what would be covered in class, again respondents tended to agree (3.00 and 3.11 among instructors and students, respectively). While slightly lower, instructors and students rated promotion of active learning in the classroom relatively highly (2.81 and 2.95, respectively), and a similar finding emerged for pre-class activities supporting increased participation (2.81 and 2.74 among instructors and students, respectively).

Table 5. Student and instructor perceptions of pre-class activities in Achieve.

Pre-class activities	Student average	Instructor average
Pre-class activities helped me stay on track with the reading	2.96	3.00
Pre-class activities helped my students stay on track with the reading		
Pre-class activities helped me achieve a basic understanding of concepts	3.08	3.03
Pre-class activities gave my students a basic understanding of concepts		
Pre-class activities gave me a basic understanding of what we would be covering in class	3.11	3.00
Pre-class activities gave my students a basic understanding of what we would be covering in class		
Pre-class activities helped me actively learn in the classroom	2.95	2.81
Pre-class activities enabled actively learn in the classroom		
Pre-class activities helped me participate more in class than I normally do	2.74	2.81
Pre-class activities promoted more classroom participation than there typically is in this course		



Discussion

Engagement in activities that were assigned to be completed before coming to the class in which that content would be covered was positively related to student performance on Achieve summative assessments and in-course final exams. We suggest that pre-class engagement predicts assessment performance because students who engage in pre-class activities get more exposure to the material than students who don't engage in pre-class activities. Also, pre-class activities provide an introduction to basic content, reducing the need for direct instruction in class and enabling more higher order thinking and opportunities for active learning during in-course time.

We also found that the more assigned pre-class activities in Achieve that a student engaged in, the higher their summative assessment scores in Achieve. Again, this could be explained in a few ways. It could be that the more pre-class activities a student was assigned, the more their instructor could focus on higher-order thinking in the classroom thus supporting increased mastery. It could also be that the more pre-class activities a student engaged in, the more exposure they had to the material, and that practice led to comprehension and improved retrieval on the summative assessment. Or, some combination of these explanations.



A subsequent analysis found that student performance on summative assessments was related to the proportion of assigned pre-class activities that they engaged in. We reasoned that the relationship might be influenced by prior academic performance, current academic performance, level of motivation to succeed, whether the student was taking the course as part of a major requirement, and their comfort level using technology. We found that level of motivation, whether the course was being taken as part of a major, and comfort level did not help the model fit better and were removed. We concluded that the proportion of assigned pre-class activities that a student engaged in significantly predicted average summative assessment score even when holding prior and current academic performance constant. Thus, we can infer that engaging in pre-class activities benefits subsequent higher stakes performance among students of all levels of preparation to succeed. We investigated the validity of the findings by exploring the incremental change in average summative assessment score by engagement score band and found that there was an obvious relationship that persisted when examined within college readiness status.

Most previous research on utilization of pre-class activities compares use with no use, but does not disaggregate the levels of utilization. This research expands the conversation by investigating both use in general, and the extent to which students utilized the tool.

Some researchers have questioned the practical significance of the research findings of educational technology tools on learning outcomes.

In the first and second research questions we found that the average difference in assessment scores would represent more than half of a grade in most higher educational institutions, and that the relationships were statistically significant regardless of prior academic performance, baseline level of motivation to succeed, and who the instructor was. It is reasonable to expect that instructors and students alike would find these differences to be practically significant, particularly when the low average performance of this sample is considered - this difference could bring students from a 65 to above a 70 on their assessment scores.

Similarly, in the third research question in this study, the correlation between engagement in pre-class activities and summative assessment scores in Achieve had a strong correlation by Cohen's (1988) standards and when disaggregated by engagement band, at utilization level $\geq 81\%$ the average summative assessment score was about six percentage points higher than the group at utilization level $\geq 61\%$ and $< 81\%$. This difference also represents about one half of a grade summative assessment average. One can assume that instructors and students would see that difference as practically significant. And, although the correlation between utilization of pre-class activities and final exam scores was weak by Cohen's standards, when disaggregated by engagement band, at utilization level $\geq 81\%$ the average final exam score was about six percentage points higher than the group at utilization level $\geq 61\%$ and $< 81\%$. Again, it is fair to assume that most students would find practical significance in this difference.



We expected that use of pre-class activities in Achieve would positively influence a student's perception of Achieve. We concluded that assigning pre-class activities positively influenced the extent to which students would recommend a course to a friend if they knew that Achieve was going to be used after accounting for comfort with technology and prior academic performance. We were slightly surprised at what a small proportion of the variability use of pre-class activities accounted for, but it is reasonable since there are many other factors that more substantially influence whether a student would recommend a course to a friend. The findings that examine differences in perception of Achieve based on group are likely more valid in this study. It was interesting to uncover that use of an additional component of Achieve did not negatively influence ease of use perceptions and that students and instructors had positive perceptions. It is reasonable to suggest that instructors might consider assigning pre-class activities and not have concerns about increasing ease of use or technological complexity burden on students.

In summary, in addition to contributing to the previous research on flipped classrooms and pre-class activities in general, this is the first research study to be conducted on Achieve, a new digital learning tool and extends the research by adding a new digital learning platform to the body of literature.

The research was conducted with a sample of instructors before being used at scale because we wanted instructors to have evidence to base their adoption and implementation decisions on immediately upon Achieve being available in the market. We suggest that the insights that emerged allow instructors to be confident that if they implement pre-class activities in Achieve in their course(s) their students will have positive perceptions of them and they will contribute to subsequent students' academic success in the course.



Limitations and Future research

Like most applied research, this study is not without limitations. First, this research is correlational so causal inferences cannot be derived from the results. To measure the causal impact of pre-class activities in Achieve an experimental design should be employed. A comparison study of Achieve which is planned for the Spring 2020 semester. It is important to first have empirical evidence of implementation patterns and the outcomes realized within those patterns before attempting to isolate impact, especially for a tool that has just been developed.

Also, engagement in a Achieve pre-class activity was defined as having launched the activity and submitting at least one response within the activity. That is, a student who launched an activity and completed one item is categorized the same way as a student who launched an activity and completed all items. One of our peer reviewers has suggested that “initiation” is a better description than “engagement” and that “engagement” should be defined by some amount of use. However, “amount” seems difficult to standardize. We can attempt to measure levels of engagement in two ways. One, by number of items attempted, but since instructors set a target score and one of the pre-class activities is adaptive, the number of items a student is exposed to varies greatly. Second, by time in the activity, but we cannot validate that time on task is a reliable proxy since we have no insight into whether a student was engaged for the duration of that time. So, for purposes of standardization we maintained that “engagement” is defined as launch and completion of at least one item in the activity. But, in future research we will seek to gain a more precise definition for engagement.



Further, there were five disciplines represented in this study. As shown in table one there seem to be differences among disciplines. Much of this we tried to account for by controlling for instructors, but current research is replicating the analyses within discipline where sample sizes allow.

Additionally, while we hypothesized that engagement in pre-class activities in Achieve would influence subsequent assessment performance, we acknowledge that the instruction provided in the classroom is of critical importance to student success. We suggest that the most effective model is one where instructors use the insights gleaned from pre-class performance to direct their in-class instruction. In this study we did not include any information on in-class instruction. In the replication study we are more systematically collecting data on whether instructors use insights gleaned from pre-class activities to influence in-class methods. Future analyses will include these data and answer the research questions that include in-class time. For example, does use of pre-class activities plus the implementation of active learning in the classroom influence assessment performance more than use of pre-class activities without implementation of active learning strategies?

Finally, Achieve was evaluated before it was used at scale in market in order to give instructors evidence for decision making about adoption and implementation once it is available for sale. As such, the instructors who participated in this study acknowledged that they were the first adopters and may have been more tolerant to a new digital learning platform that was still being optimized. Currently, a replication study is being conducted with an instructor sample more representative of the population of instructors to establish whether these findings persist and are therefore more generalizable.



Note on data privacy

Prior to data collection, this study and the associated consent forms and instruments were reviewed and approved (found exempt) by the Human Resources Research Organization (HumRRO). HumRRO is a third-party Institutional Review Board organization with no affiliation with Macmillan Learning (federal wide assurance number 00009492 and IRB number 00000257). Macmillan Learning seeks independent and unfunded third-party review to eliminate any bias in decision of exemption. Macmillan Learning then seeks local Institutional Review Board approval at each participating institution, where required. The data collected in this study, which are provided by the instructor and consenting students, are initially identifiable. However, once a random identifier is generated identifiable data are destroyed. Data are provided in secure storage locations, and access is permitted only to the primary investigator in the study. For full details of our data handling and storage privacy procedures, contact Kara McWilliams, Vice President Impact Research at Macmillan Learning at kara.mcwilliams@macmillan.com.



REFERENCES

- Abeyskera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development*, 34(1), 1-14. Retrieved from <http://dx.doi.org/10.1080/07294360.2014.934336>
- Albalawi, A.S. (2018). The effect of using flipped classroom in teaching calculus on students' achievement at University of Tabuk. *International Journal of Research in Education and Science (IJRES)*, 4(1), 198-207. Retrieved from <https://www.ijres.net/index.php/ijres/article/view/187/pdf>
- Baepler, P. M., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers and Education*, 78, 227-236. Retrieved from <https://doi.org/10.1016/j.compedu.2014.06.006>
- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. Proceedings of the First International Conference on gameful design, research, and applications, 10-17. Retrieved from https://www.researchgate.net/publication/259821680_Improving_Participation_and_Learning_with_Gamification
- Basal, A. (2015). The implementation of a flipped classroom in foreign language teaching. *Turkish Online Journal of Distance Education*, 16(4), 28-37. DOI: 10.17718/tojde.72185
- Bishop, J., & Verleger, M. A. (2013, June). The flipped classroom: A survey of the research paper presented at 2013 ASEE Annual Conference & Exposition, Atlanta, Georgia. Retrieved from <https://www.asee.org/public/conferences/20/papers/6219/view>
- Brown, P. C., Roediger, H. L. III, & McDaniel, M. A. (2014). Make it stick: The science of successful learning. Cambridge, MA, US: Belknap Press of Harvard University Press. Retrieved from <https://www.hup.harvard.edu/catalog.php?isbn=9780674729018>
- Cabi, E. (2018) The Impact of the Flipped Classroom Model on Students' Academic Achievement. *International Review of Research in Open and Distributed Learning*, 19. Retrieved from <https://doi.org/10.19173/irrodl.v19i3.3482>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers. Retrieved from [https://www.scrip.org/\(S\(z5mqp453edsnp55rrgct55\)\)/reference/ReferencesPapers.aspx?ReferencID=2041144](https://www.scrip.org/(S(z5mqp453edsnp55rrgct55))/reference/ReferencesPapers.aspx?ReferencID=2041144)
- Cummins, S., Beresford, A., & Rice, Andrew. (2015). Investigating Engagement with In-Video Quiz Questions in a Programming Course. *IEEE Transactions on Learning Technologies*, 9, 1-1. Retrieved from https://www.researchgate.net/publication/279231499_Investigating_Engagement_with_In-Video_Quiz_Questions_in_a_Programming_Course
- Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development*, 61(4), 563-580. Retrieved from <https://link.springer.com/article/10.1007/s11423-013-9305-6>
- Dicheva, D., & Dichev, C. (2016). An Active Learning Model Employing Flipped Learning and Gamification Strategies. In Proceedings of first Int. Workshop on Intelligent Mentoring Systems (IMS 2016), Zagreb, Croatia (pp. 1-6). Retrieved from https://www.researchgate.net/publication/301961326_An_Active_Learning_Model_Employing_Flipped_Learning_and_Gamification_Strategies
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8410-8415. Retrieved from https://www.researchgate.net/publication/301961326_An_Active_Learning_Model_Employing_Flipped_Learning_and_Gamification_Strategies
- Freudenthal, H. (1983) Didactical Phenomenology of Mathematical Structures (Dordrecht, The Netherlands: Reidel). Retrieved from http://gpdmatematica.org.ar/wp-content/uploads/2015/08/Freudenthal_Didactical_Phenomenology_of_Mathematical_Structures1983.pdf
- Graziano, K. J. (2017). Peer teaching in a flipped teacher education classroom. *TechTrends*, 61(2), 121-129. Retrieved from https://www.researchgate.net/publication/302922870_Peer_Teaching_in_a_Flipped_Teacher_Education_Classroom
- Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A Measure of College Student Course Engagement. *The Journal of Educational Research*, 98(3), 184-191. Retrieved from <http://dx.doi.org/10.3200/JOER.98.3.184-192>
- Kim C. Park, & S. W. Cozart J. (2014). Affective and motivational factors of learning in online mathematics courses. *British Journal of Educational Technology*, 45(1), 171-185. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8535.2012.01382.x>
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature. *Review of Educational Research*, 75(1), 63-82. Retrieved from <https://doi.org/10.3102/00346543075001063>
- Latulipe, C., Long, N. B., & Seminario, C. E. (2015). Structuring flipped classes with lightweight teams and gamification. SIGCSE, 2015, Proceedings of the 46th ACM Technical Symposium on Computer Science Education (pp. 392-397). New York, NY: ACM. Retrieved from https://www.researchgate.net/publication/297663593_Structuring_Flipped_Classes_with_Lightweight_Teams_and_Gamification
- Pierce, R., & Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *American journal of pharmaceutical education*, 76(10), 196. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530058/>
- Shaw, E. J., & Mattern, K. D. (2009). Examining the accuracy of self-reported high school grade point average. (College Board Research Report No. 2009-5). New York: The College Board. Retrieved from <https://eric.ed.gov/?id=ED562616>
- Smallhorn, M. (2017). The flipped classroom: A learning model to increase student engagement not academic achievement. *Student Success*, 8(2). Retrieved from <https://studentsuccessjournal.org/article/download/502/368>
- Sun, J. C.-Y., & Wu, Y.-T. (2016). Analysis of Learning Achievement and Teacher-Student Interactions in Flipped and Conventional Classrooms. *The International Review of Research in Open and Distributed Learning*, 17(1). Retrieved from <https://doi.org/10.19173/irrodl.v17i1.2116>
- Tawfik, A.A., & Lilly, C. Using a flipped classroom approach to support problem-based learning. *Technol. Knowl. Learn.* 2015, 20, 299-315. Retrieved from https://www.researchgate.net/publication/282553658_Using_a_Flipped_Classroom_Approach_to_Support_Problem-Based_Learning
- Wiginton, B.L. (2013). Flipped Instruction: An Investigation into the Effect of Learning Environment on Student Self-Efficacy, Learning Style, and Academic Achievement in an Algebra I Classroom. (Doctoral dissertation). The University of Alabama. Retrieved from <https://search.proquest.com/docview/1505373684?pq-origsite=gscholar>
- Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior*, Vol. 70, pp. 251-260. Retrieved from <https://dl.acm.org/citation.cfm?id=3062022>
- Zainuddin, Z. (2018). Students' learning performance and perceived motivation in gamified flipped-class instruction. *Computers & Education*, 126, 75-88. Retrieved from <https://www.deepdyve.com/lp/elsevier/students-learning-performance-and-perceived-motivation-in-gamified-NkOBGMkzKF>
- Zengin, Y. (2017). Investigating the use of the Khan Academy and mathematics software with a flipped classroom approach in mathematics teaching. *Journal of Educational Technology and Society*, Vol. 20 No. 2, pp. 89-100. Retrieved from https://www.researchgate.net/publication/315696761_Investigating_the_Use_of_the_Khan_Academy_and_Mathematics_Software_with_a_Flipped_Classroom_Approach_in_Mathematics_Teaching

About the Authors

Dr. Kara McWilliams

Vice President, Impact Research

Kara is passionate about researching the impact of digital technologies in higher education, and how insights can inform teaching and learning. She has ten years of experience conducting qualitative and quantitative investigations of how course and classroom interventions can improve learner outcomes and influence learning gains. She holds a doctorate in Educational Research, Measurement and Evaluation and a master's degree in Curriculum & Instruction from Boston College.

Dr. Jeff Bergin

Vice President, Learning Research & Design

Jeff has led curriculum, instructional, and learner experience design for various educational technology companies for the past 20 years. Dr. Bergin leads a team of learning researchers and experience designers. In previous roles, he has led the design of personalized and mobile products. Dr. Bergin holds a Ph.D. from Arizona State University and has presented and published on topics including learning design, online learning, and technology-augmented instruction.

Alexandra Stone, University of Connecticut

Alexandra Stone, University of Connecticut. Alexandra is a graduate student at the University of Connecticut working on her Ph.D. in Measurement, Evaluation, and Assessment where she earned her M.A. in the same program. At the time of this study Alexandra was an Associate Impact Research Scientist in the Learning Science and Insights team at Macmillan Learning, she continues to contract for the Learning Science team. Alexandra has broad experience in using qualitative and quantitative methods to examine the effectiveness of instructional programs on student achievement. She also has experience in developing and managing large databases using SQL and performing data analyses in SAS and R. Alexandra received a B.A. in Statistics from Elon University.

About Macmillan Learning

Macmillan Learning improves lives through learning. Our legacy of excellence in education continues to inform our approach to developing world-class content with pioneering, interactive tools. Through deep partnership with the world's best researchers, educators, administrators, and developers, we facilitate teaching and learning opportunities that spark student engagement and improve outcomes. We provide educators with tailored solutions designed to inspire curiosity and measure progress. Our commitment to teaching and discovery upholds our mission to improve lives through learning. To learn more, please visit <http://www.macmillanlearning.com> or see us on Facebook, Twitter, LinkedIn or join our Macmillan Community.

About the Learning Science and Insights Team

As the Learning Insights company, we are passionate and scientific about helping students, instructors, and institutions to achieve their full potential. We use a unique combination of user-centered design, research from the learning sciences, and empirical insights from extensive data mining and Impact Research. To learn more about this approach, please visit <http://www.macmillanlearning.com/catalog/page/learningscience>