*Research Article*

*Educational Evaluation and Policy Analysis*
*December 2021, Vol. 43, No. 4, pp. 668–687*
*DOI: 10.3102/01623737211013782*
*Article reuse guidelines: sagepub.com/journals-permissions*


# Mobilizing Developmental Education: The Causal Effect of Mobile App Courseware on the College Outcomes of Developmental Education Students

**Matt S. Giani**
*The University of Texas at Austin*

**Allison Martin**
*Bossier Parish Community College*

*Developmental education, in which college students deemed unprepared for college-level coursework enroll in non-credit-bearing courses, is widespread in American higher education. This study evaluates the effect of mobile app courseware on the college outcomes of developmental education students. We used a research design that randomly assigned course sections to receive access to the apps or not. The results show that access to the apps significantly improved student performance in developmental education outcomes and marginally improved medium-term college persistence and performance but did not affect credential attainment in the study timeframe. Despite a number of barriers to implementation, the results suggest the intervention has the potential to improve the short-term outcomes of developmental education students in addition to being low-cost and scalable.*

Keywords: *community colleges, evaluation, postsecondary education, higher education, experimental research, focus group interviews*

Developmental education, in which college students deemed unprepared for college-level coursework enroll in non-credit-bearing courses to improve their academic skills, is widespread in American higher education. Among students in the United States who began postsecondary education in 2011–2012, 65% of those starting at public 2-year colleges and 31% of those beginning at public 4-year colleges enrolled in at least one developmental education course over the next 6 years (X. Chen et al., 2020). Yet the majority of students assigned to developmental education never complete the required sequence of developmental courses (Bailey et al., 2008), and less than a quarter of community college students who take developmental courses earn a degree from any institution within 6 years (Ganga et al., 2018). Although some research suggests that the outcomes of developmental education students are no different from community college students who do not enroll in developmental education once other student characteristics are controlled for (Bahr, 2008), reviews of the developmental education literature suggest placement into developmental education may have a significant negative effect on students' probability of passing the college-level course in the developmental subject, college credits earned, and attainment (Valentine et al., 2017).

Due to the prevalence of developmental education and the sub-optimal outcomes of students assigned to remediation (developmental education and remediation are used interchangeably throughout), researchers and reformers have developed a number of strategies to improve developmental students' performance, persistence, and attainment (Bettinger et al., 2013). Reforms such as integrating developmental content into other academic and technical courses

(Wachen et al., 2012), providing greater academic and financial supports to developmental education students (Scrivener et al., 2015), and creating learning communities for academically underprepared students (Sommo et al., 2012; Visher et al., 2012) can improve students' performance in developmental courses and increase persistence and attainment rates. However, many of these reforms are also quite expensive, raising questions about their efficiency and sustainability (Bettinger et al., 2013).

Bossier Parish Community College (BPCC), a mid-sized community college near Shreveport, Louisiana, developed a resource called Open Campus™ that contains free and fully online modules of developmental education content. Open Campus was primarily designed to help students brush up on core academic topics before taking developmental education placement exams, but the progenitors of Open Campus soon realized that faculty were using the resource in their instruction to allow students to review and practice concepts covered in class with which they had struggled. However, BPCC administrators also learned that accessing Open Campus was an issue for many students, particularly those from low socioeconomic households, who did not have a computer and/or Internet access at home.

In 2015, BPCC received a grant from the U.S. Department of Education through the First in the World (FITW) program. To further increase access to Open Campus content, BPCC proposed in its FITW application to convert Open Campus into a mobile application for three of its most common developmental courses: Math 98 (Beginning Algebra I), Math 99 (Beginning Algebra II), and English 99 (Fundamentals of Composition). BPCC also proposed to evaluate the effect of access to these mobile applications on student outcomes through a randomized block design (RBD). Developmental education faculty served as blocks, and their course sections were randomly assigned to receive access to the mobile apps (the treatment condition) or not (the control condition). Two pilots were administered, one in Spring 2017 and another in Fall 2017, with roughly 2,000 students participating in the pilots. The purpose of this study is describe the impact of access to mobile applications on students' performance in developmental education and subsequent courses, retention at BPCC, and attainment of credentials through Spring 2019.

## Literature Review

Given the scope of student participation in developmental education courses nationally and the discouraging success rates of students assigned to developmental education, a broad range of strategies have been developed to improve developmental education student outcomes. These reforms can be placed into one of three broad categories: changing how and whether students are assigned to developmental education, pedagogical and curricular reforms, and supplemental services and supports. Although a number of thorough reviews now exist cataloging these reforms (Edgecombe et al., 2013) and assessing the evidence base on these strategies (Bailey et al., 2017; Bettinger et al., 2013; Zachry & Schneider, 2011), they are briefly reviewed here to contextualize the current reform.

### *Developmental Education Placement Reforms*

Some studies have shown as many as one third of students are severely misplaced when comparing their scores on developmental education placement exams with their performance in subsequent college courses (Scott-Clayton, 2012). Concerns about misplacement, and the overreliance on standardized assessments to place students into developmental education (Hughes & Scott-Clayton, 2011; Sedlacek, 2004), has led to growing support for the use of multiple measures to make developmental education placement decisions. While the evidence base on the effect of implementing multiple measures, such as high school grade point average (GPA), completion of specific high school courses, and noncognitive assessments (Bailey et al., 2017), is thin, preliminary evidence from recent studies suggests the use of multiple measures may significantly increase the rates that students are placed directly into college-level courses and pass those courses (Barnett et al., 2018).

Another approach to reducing the number of students enrolling in developmental education is to improve students' readiness for college-level coursework before they graduate high school

(Fay et al., 2017). A number of states have experimented with testing students in Grade 11 to determine their college readiness and, in the cases of Florida, Tennessee, and West Virginia, assigning them to college readiness courses in Grade 12 if they scored below the threshold set for college readiness. Research from California (Jackson & Kurlaender, 2016; Kurlaender et al., 2020), Florida (Mokher et al., 2018), Maine (Hurwitz et al., 2015), Tennessee (Kane et al., 2021), and West Virginia (Pheatt et al., 2016) has found mixed but somewhat positive results of these types of policies on improving college-going rates and reducing developmental education enrollments, although limited research exists on the long-term effects of these policies.

### *Curricular and Pedagogical Reforms*

Although the reforms discussed above may result in more accurate placements into developmental education, some students may still be unprepared for college-level coursework even if placements are more accurate. Curricular and pedagogical reforms are therefore being devised and tested to determine whether they improve developmental education student outcomes. The most notable examples of this strategy are acceleration, contextualization, and co-requisite approaches.

As community colleges often require students to complete developmental education courses before enrolling in credit-bearing courses, acceleration strategies attempt to reduce the time it takes for students to complete the developmental sequence (Edgecombe et al., 2013; Nodine et al., 2013; Venezia & Hughes, 2013). This is often done by compressing courses into half-semester or shorter periods, allowing students to potentially complete multiple developmental education courses in a single semester. Experimental research on the efficacy of this strategy is limited (Venezia & Hughes, 2013), but correlational and quasi-experimental studies provide suggestive evidence that accelerating developmental education courses may promote students' entry into college-level coursework (Bailey & Jaggars, 2016; Fay et al., 2017; Hodara & Jaggars, 2014; Jaggars et al., 2015).

Contextualized models integrate developmental education content into what are traditionally nondevelopmental education courses. One of the first examples of this approach was the integrated basic education and skills (I-BEST) model developed in Washington state (Wachen et al., 2012; Zeidenberg et al., 2010). Though targeted primarily to adult workers in career-oriented fields, the idea behind I-BEST was to enroll students directly into career-focused courses but integrate developmental education content into those courses. A critical ingredient in the I-BEST model was the courses were co-taught by a basic skills instructor and an occupational instructor (Zeidenberg et al., 2010). Quasi-experimental evidence suggests the program significantly improved students' likelihood of completing credentials (Wachen et al., 2012).

The contextualized approach led to a related strategy of co-requisite remediation (also known as mainstreaming), in which students who would have traditionally been required to complete developmental education before enrolling in credit-bearing courses are allowed to enroll directly in credit-bearing courses. However, the key innovation is students are required to enroll in developmental education courses or complete supplemental developmental education content (through labs or "boot camps") simultaneously. The expansion of co-requisite remediation has been promoted by organizations such as Complete College America, and today approximately 20 states allow or mandate co-requisite approaches to remediation (Education Commission of the States, 2019). Although rigorous empirical evidence on this strategy is just emerging, at least one experimental study has found that being assigned to a co-requisite math course significantly improves long-term persistence and attainment when compared with traditional algebra remediation (Logue et al., 2019).

### *Supplemental Services and Supports*

Some have argued that even greater reforms are needed in developmental education to truly move the needle on students' long-term persistence and attainment outcomes in a scalable and sustainable way (Kane et al., 2021). Community colleges have begun to experiment with substantive reforms aimed at providing a suite of wraparound supports to students in developmental education courses. The most widely known of these strategies is the City University of New

York's (CUNY) Accelerated Study in Associate Programs (ASAP) model. Although not targeted at developmental education students initially, the program's eligibility criteria expanded to include developmental education students a few years after its inception.

The ASAP model comprises a host of integrated reforms aimed at increasing students' academic momentum and sense of belonging at the college (Boykin & Prince, 2015). These reforms include the requirement for students to be enrolled full-time each semester, a cohort model where students take classes with students in the same major, financial incentives (such as tuition waivers, textbook assistance, and MetroCards), dedicated advisors and mandatory advisement, and mandatory tutoring specifically for students placed into developmental education courses. CUNY partnered with MDRC to evaluate the efficacy of the program using a randomized controlled trial design, and the results show that the ASAP treatment group's 3-year graduation rate was almost double that of the control group (40% vs. 22%; Scrivener et al., 2015). An evaluation of a replication of the ASAP model in other Ohio has found similarly pronounced effects (Sommo et al., 2018), with future evaluations of other replications forthcoming.

## Conceptual Framework—Learning Science Principles, Mobile Learning, and Developmental Education

The research discussed above has found that the persistence and attainment rates of students who begin in developmental education can be substantially improved. However, in an era of declining appropriations for higher education in many states, the sustainability and scalability of reforms that often require significant recurring investment of resources, such as co-taught courses (Belfied et al., 2016) and wraparound services and supports (Scrivener et al., 2015), are a concern. For example, despite ASAP's substantial effect on college graduation, the New York City mayor's office announced a US$20 million cut to CUNY ASAP and its suspension in 2020–2021 (St. Amour, 2020), and Ohio's replication of ASAP was not sustained in two of the three colleges that participated in the evaluation (Miller et al., 2020). Particularly in the harsh financial realities emerging due to the COVID-19 pandemic and the economic recession it is causing, many institutions are searching for low-cost alternatives to move the needle on developmental education student outcomes.

One such alternative is mobile-enabled learning or "m-learning," which is on the rise in both informal and formal contexts in U.S. higher education (Alrasheedi et al., 2015; B. Chen & deNoyelles, 2013; Crompton & Burke, 2018; Kaliisa & Picard, 2017; Pimmer et al., 2016). There are four primary reasons m-learning has been theorized to be an effective strategy for promoting achievement and attainment in higher education. First, m-learning is one of many strategies and technologies that can be used to promote college students' self-regulated learning, an important consideration for students placed into developmental education (Bailey et al., 2017; Young & Ley, 2003). Specifically, students enrolling in college for the first time may not have a clear sense of their preparedness for college-level coursework or the specific gaps in their knowledge that may hinder their understanding of more complex topics. M-learning can include diagnostic exams that provide students a clearer sense of the content areas in which they should prioritize their learning and personalized learning paths based on the gaps in their knowledge, a mobile-enabled example of the technique historically referred to as "programmed instruction" (Kulik et al., 1980; Skinner, 1954; Thorndike, 1912). Traditionally, few colleges use a self-placement model, where students use the results of assessments to choose their own learning path (Hodara et al., 2012). M-learning also facilitates the use of practice quizzes that students can take and retake, a strategy known as "retrieval practice" that has shown to be strongly related to long-term retention (Roediger & Butler, 2011).

Second, m-learning can enable hybrid, blended, or "flipped" classroom environments, in which students use technological resources to learn some of the content outside of the class and then apply those concepts to examples in class, alongside their peers and under the guidance of an in-person instructor (Bersin, 2004; Garrison & Kanuka, 2004; Graham et al., 2013; Thorne, 2003). This approach is referred to as "flipped learning" as it inverts the traditional pedagogical approach of the instructor lecturing in class and students completing assignments on their own outside of class.

Third, technological resources also allow for the "gamification" of learning, which has been shown to increase student engagement and academic performance (Kapp, 2012). Gamification comprises three key components. First, content is broken down into constituent elements, known as "modularization." For example, students are able to maintain engagement more easily when viewing a series of 3-minute videos on specific topics rather than an entire 45-minute lecture covering a range of topics. Second, students receive immediate feedback and recognition for their accomplishments through points, badges, or "leveling up." This allows students to recognize the progress they are making, which motivates persistence and engagement. Third, gamified learning experiences are student-centered, where the learner has the autonomy to choose the path and pace that is best for them. Classroom environments are often teacher-centered and structured around a specific syllabus common to all students.

Fourth, quite simply, mobile devices are nearly ubiquitous. Roughly 95% of the global population live in an area with cellular coverage, and 84% have access to mobile-broadband networks (3G or above; International Telecommunications Union, 2016). In the United States, young adults aged 18 to 29 years are the demographic with the highest ownership rate of mobile devices, and young adults, low-income adults, people of color, and adults with lower levels of education are more likely to be "smartphone dependent"—those who access the internet from a mobile device but do not own a laptop or desktop computer (Pew Research Center, 2019). Although the benefits of m-learning discussed above may also be realized by other technologies, such as desktop or laptop computers, m-learning may be a particularly important strategy for community colleges that tend to serve high proportions of students from populations with greater access to mobile devices than computers.

Despite the promise of m-learning, many studies have concluded that high-need students struggle in purely online environments, but are motivated to persist when the technology is a function of how they engage outside the formal classroom (Xu & Jaggars, 2011). An important ingredient in student engagement is whether the student perceives a connection of any kind to the campus, knows an instructor or student's name, or feels someone knows theirs (Center for Community College Student Engagement, 2013). The question for colleges is therefore how to realize the benefits of incorporating m-learning into instruction without using it to replace the bonds and relationships more easily cultivated by in-person environments.

A number of reviews now exist summarizing the evidence base on m-learning in higher education (Alrasheedi et al., 2015; B. Chen & deNoyelles, 2013; Crompton & Burke, 2018; Kaliisa & Picard, 2017; Pimmer et al., 2016). However, the most recent review estimates that only six empirical studies used an experimental design to estimate the effect of m-learning on student outcomes (Crompton & Burke, 2018). To the authors' knowledge, no studies have used an experimental design to examine the efficacy of this strategy on students placed into developmental education courses specifically, making this study the first such evaluation.

## Open Campus™ Mobile App Intervention

While the Open Campus mobile apps are currently available to download through the Apple App Store and Google Play (screenshots of the mobile apps are included in supplemental Figure 1 in the online version of the journal), a beta version of the apps was evaluated in this study. This version of the intervention would be more accurately described as a mobile-responsive website that students were instructed to save as an icon on the home screen of their mobile phone. After doing so, the resource functioned very similarly to a native mobile app, and we refer to the intervention as a mobile app throughout for simplicity. However, the implementation evaluation found that this format of the "apps" was confusing to both students and instructors, potentially hindering students' use of the apps. We will return to these points in our discussion.

Once students download the app and create an account, they are presented with various courses that they may choose from. For the present evaluation, students could enroll in English 99 (Developmental Writing), Math 98 (Beginning Algebra I), or Math 99 (Beginning Algebra II).

After enrolling in any of the above courses, students take a preassessment within the apps with roughly 50 items that further diagnoses the specific areas in which they need the greatest assistance. After completing the quiz, students are shown the percentage of questions that they got right and are instructed to focus their attention on the modules that are aligned with the topics with which they struggled the most. For example, the Math 98 course is divided into 22 separate modules or topics, and those with which the student struggled would become highlighted with a gold star based on the preassessment results.

After the preassessment, students can select whatever module they prefer among those modules starred for their completion. Each module contains a separate 10-item quiz with additional questions related to that specific topic area. For the module quizzes, students are shown whether they got each question correct or incorrect. For each incorrect answer, they are told the specific module they should review to learn more about that topic. For example, in Math 98's Module 5 on "Solving One Step Equations With Addition or Multiplication," the student may be instructed to "please go back and watch the video for Module 5.1 on equation vocabulary for addition and multiplication" if they made an incorrect choice on the quiz item aligned with that topic. An important element of within-module quizzes is that they are designed for mastery learning rather than evaluation—students are able to attempt the quizzes as many times as they would like, and only their highest grade is kept in their profile.

Each module also contains a series of video lectures where BPCC instructors explain the concepts. Although the desktop version of Open Campus was widely used, BPCC administrators noticed through analytics tracking that many users would discontinue watching the Open Campus videos after 3 to 5 minutes, despite the videos being 20 to 30 minutes long on average. During the conversion of Open Campus from desktop to mobile, the videos were further modularized so that each video is now only 3 to 6 minutes long.

In addition to the core pedagogical features of quizzes and videos included in each course, the mobile apps contain supplemental features meant to increase engagement. First, students can earn digital badges based on their progress within the modules that also confer elements of attire for a digital avatar ("My Cavalier"). This avatar was designed by computer science students at BPCC and is based on BPCC's mascot. Second, students can also challenge each other, which lets them go head-to-head in taking a quiz. The apps contain a leaderboard showing the users with the most wins, both overall and for specific courses. Third, the app sends email nudges to students if they have been inactive in the app for a prolonged period of time or if they have earned badges signifying major milestones in the app (e.g., 50% of the module has been completed). As mentioned above, the email notifications are an opt-in feature, and students can unsubscribe from the notification in-app or through email.

## Method

### *Research Questions*

The study address three research questions:

**Research Question 1:** What is the effect of access to mobile applications on the postsecondary outcomes of developmental education students?

**Research Question 2:** To what extent does this effect vary across courses (English 99, Math 98, and Math 99)?

**Research Question 3:** What is the effect of the use of the mobile applications on the postsecondary outcomes of developmental education students?

The first and second research questions call for an *intent-to-treat* (ITT) analysis (Gupta, 2011), where the outcomes for the entire sample of students assigned to the treatment are compared with the outcomes of the entire sample of students assigned to the control group. When a randomized controlled trial is used for this type of analysis, as is the case in this study, it can be referred to as an *effectiveness trial*. These analyses produce estimates of the effect of the treatment in routine conditions. However, there is also theoretical and policy interest in the effect of the intervention for students who likely use the mobile apps, also known as the effect of the *treatment-on-the-treated* (TOT) or the *complier average causal effect* (CACE).

For example, if technological barriers or other issues prevented the entire treatment group from using the mobile apps and those barriers could be removed in future implementations, the TOT estimates allow us to project what the effects may be with increased take-up of the intervention. The sections below describe our research design and statistical approach that allows us to produce both ITT and TOT estimates.

### *Research Design*

Students were randomly assigned to either the treatment or the control group using a design known as a randomized block design or stratified cluster randomized controlled trial (Bloom et al., 1999; Hedges & Hedberg, 2007; Murray, 1998; Murray et al., 2004). We elected to use a group random assignment rather than individual assignment given the threat of "contamination," namely, students who would be given access to the mobile apps sharing with control students enrolled in the same developmental course sections. In addition, faculty members expressed a strong preference for cluster randomization to eliminate the need to assist students in the same classrooms using different materials and technologies.

The four strata used in the randomization are college (BPCC or Northwestern State University in Louisiana [NSU]), semester (Spring 2017 or Fall 2017), developmental education course (Math 98, Math 99, and English 99), and instructors. A block was created for every instructor who taught more than one course section of the same course in the same semester at the same college, and that instructor's specific course sections were randomly assigned to the two groups. For example, there were four instructors who each taught two course sections of English 99 during the Spring 2017 semester, and each of those instructors had one section randomly assigned to the treatment group and one assigned to the control group. An additional block was created for each college, semester, and course combination for all remaining instructors who taught one course section. For example, two instructors taught only one section of English 99 during the Spring 2017 semesters. Those two instructors were placed into the same block, and one of their course sections was placed in the treatment group, whereas the other was placed in the control group. The majority of course sections were assigned within instructor; of the total 80 course sections in the analytic sample, 48 (60%) were assigned within instructors who taught multiple sections, whereas 32 (40%) were assigned within multiple-instructor blocks comprised of instructors only teaching one section of that course. The design produced 17 blocks during the Spring 2017 intervention and 14 blocks during the Fall 2017 intervention for a total of 31 blocks, each of which contained treatment and control sections. These blocks, the number of students assigned to the treatment and control groups within each block and overall, and whether the randomization occurred within instructor (i.e., for an instructor teaching multiple sections of the same course) are reflected in Table 1.

### *Sample*

The sample is drawn from BPCC, a community college located in the Shreveport metropolitan area in northwest Louisiana. Roughly 35% of students in the analytic sample were first-time BPCC students at the time of the intervention, whereas the remaining 65% had attempted at least one course at BPCC prior to the intervention. For the latter group of returning students, the median number of credits attempted prior to the intervention was 24. The analytic sample was 64% female and 36% male. Approximately 65% of the sample received a Pell grant, with a mean Pell award of US$2,442 in the semester of the intervention for students who received a Pell grant. The mean age of the sample was 25 years. Although the race/ethnicity data provided for the study did not appear to be reliable (more than half of students had "other" indicated for their race/ethnicity), the BPCC population as a whole is roughly 46% White non-Hispanic, 40% Black, and 6% Hispanic/Latinx, with all other racial/ethnic groups comprising the remaining 8%.

BPCC piloted the intervention in two semesters, Spring 2017 (Cohort 1) and Fall 2017 (Cohort 2). For the Spring 2017 cohort, the total number of students who were enrolled in one of the targeted courses at the time of randomization included 866 unique students comprising 980 total course enrollments. Of the 866 unique students, 760 only enrolled in one of the three

TABLE 1

*Student Enrollment by Randomization Block for Both Cohorts*

| Block ID | Cohort | BPCC/ NSU | Course | Instructor | Control | Treatment | Total | Within |
|---|---|---|---|---|---|---|---|---|
| 1 | Spring 2017 | BPCC | English 99 | English 99 A | 18 | 23 | 41 | Yes |
| 2 | Spring 2017 | BPCC | English 99 | English 99 B | 30 | 29 | 59 | Yes |
| 3 | Spring 2017 | BPCC | English 99 | English 99 C | 18 | 27 | 45 | Yes |
| 4 | Spring 2017 | BPCC | English 99 | English 99 D | 22 | 24 | 46 | Yes |
| 5 | Spring 2017 | BPCC | English 99 | English 99 E | 28 | 26 | 54 | No |
| 6 | Spring 2017 | NSU | English 99 | English 99 F | 14 | 19 | 33 | No |
| 7 | Spring 2017 | BPCC | Math 98 | Math 98 A | 29 | 24 | 53 | Yes |
| 8 | Spring 2017 | BPCC | Math 98 | Math 98 B | 37 | 32 | 69 | Yes |
| 9 | Spring 2017 | BPCC | Math 98 | Math 98 C | 29 | 27 | 56 | Yes |
| 10 | Spring 2017 | BPCC | Math 98 | Math 98 D | 24 | 29 | 53 | Yes |
| 11 | Spring 2017 | BPCC | Math 98 | Math 98 E | 23 | 32 | 55 | Yes |
| 12 | Spring 2017 | BPCC | Math 98 | Math 98 F | 56 | 30 | 86 | No |
| 13 | Spring 2017 | NSU | Math 98 | Math 98 G | 13 | 16 | 29 | No |
| 14 | Spring 2017 | BPCC | Math 99 | Math 99 A | 34 | 28 | 62 | Yes |
| 15 | Spring 2017 | BPCC | Math 99 | Math 99 B | 33 | 33 | 66 | Yes |
| 16 | Spring 2017 | BPCC | Math 99 | Math 99 C | 87 | 56 | 143 | No |
| 17 | Spring 2017 | NSU | Math 99 | Math 99 D | 22 | 8 | 30 | No |
| 18 | Fall 2017 | BPCC | English 99 | English 99 G | 31 | 31 | 62 | Yes |
| 19 | Fall 2017 | BPCC | English 99 | English 99 H | 31 | 31 | 62 | Yes |
| 20 | Fall 2017 | BPCC | English 99 | English 99 I | 31 | 32 | 63 | Yes |
| 21 | Fall 2017 | BPCC | English 99 | English 99 J | 31 | 20 | 51 | Yes |
| 22 | Fall 2017 | BPCC | English 99 | English 99 K | 30 | 62 | 92 | No |
| 23 | Fall 2017 | BPCC | Math 98 | Math 98 H | 35 | 35 | 70 | Yes |
| 24 | Fall 2017 | BPCC | Math 98 | Math 98 I | 35 | 35 | 70 | Yes |
| 25 | Fall 2017 | BPCC | Math 98 | Math 98 J | 31 | 30 | 61 | Yes |
| 26 | Fall 2017 | BPCC | Math 98 | Math 98 K | 34 | 11 | 45 | Yes |
| 27 | Fall 2017 | BPCC | Math 98 | Math 98 L | 35 | 22 | 57 | Yes |
| 28 | Fall 2017 | BPCC | Math 98 | Math 98 M | 57 | 67 | 124 | No |
| 29 | Fall 2017 | BPCC | Math 99 | Math 99 E | 72 | 35 | 107 | Yes |
| 30 | Fall 2017 | BPCC | Math 99 | Math 99 F | 35 | 35 | 70 | Yes |
| 31 | Fall 2017 | BPCC | Math 99 | Math 99 G | 79 | 129 | 208 | No |
| Total | | | | | 1084 | 1038 | 2122 | |

*Note.* The “Within” column indicates whether the randomization of course sections occurred within instructor. Within-instructor randomization occurred for all instructors teaching multiple sections of the same course in the same semester. If “Within” is “No,” that block includes all teachers of that course in that semester who only taught one section of the course. BPCC = Bossier Parish Community College; NSU = Northwestern State University in Louisiana.

courses, 98 enrolled in two courses, and eight enrolled in all three courses. For the Fall 2017 cohort, the total number of students who were enrolled in one of the targeted courses at the time of randomization included 1,051 unique students comprising 1,142 total course enrollments. Of the 1,051 unique students, 960 only enrolled in one of the three courses and 91 enrolled in two courses. No student enrolled in all three courses in Fall 2017.

Although randomization occurred after the last date for course changes for the Spring 2017 intervention, randomization took place slightly earlier in the Fall 2017 semester to give instructors more time to prepare for using the apps in their courses. While only five of the 980 student course records did not have credit or grade information for their assigned developmental education course at the end of the semester during Spring 2017 (attrition rate = 0.5%), for the Fall

2017 semester 112 of the 1,142 course records did not have a grade (attrition rate = 9.8%). The attrition rate for the control group was 10.9% and the rate for the treatment group was 8.7%, for a differential attrition of 2.2%. Neither the overall or differential attrition rate poses a threat to the validity of the study under the What Works Clearinghouse's conservative attrition standard. Course records without a grade will be considered censored and excluded from the grade analyses (Puma et al., 2009). However, for the analysis of whether students passed the course, we consider students who withdrew or otherwise did not receive a grade as having not passed the course, and these records are kept in the analytic sample.

The analytic sample used in each model depends upon the research question being addressed. For models of outcomes that are specific to the individual developmental education courses, the entire sample was used given the limited threat of contamination. For example, it is unlikely that a student getting access to the mobile app for Math 98 would substantively affect her performance in English 99, even if she was assigned to a control group section in English. However, for outcomes related to overall course performance and long-term outcomes, an additional inclusion criterion is applied requiring students to have only enrolled in one of the three targeted course sections in that semester to ensure students are not part of both the treatment and control group. Note that the eligibility criterion (that the student enrolled in only one course) for inclusion in this sample is based on a student behavior that occurred prior to randomization, maintaining the integrity of the random assignment. The delimited sample includes 869 (50.5%) control students and 851 (49.5%) treatment students for a combined sample of 1,720 students.

### *Variables*

The outcome variables investigated in the study include (a) whether students passed the developmental education course in which they enrolled, (b) the grade they received in that course (on a 0.00- to 4.00-grade-points scale), (c) whether students persisted to the next semester, (d) the number of semester credit hours they earned in the next semester, and (e) whether they earned a credential by Spring 2019. For the attainment outcome, one model investigated whether students earned any credential, and separate models investigated whether students earned associate's degrees or certificates specifically. It should be mentioned that the time frame for investigating attainment was rather limited. Students in Cohort 1 were enrolled in developmental education courses in Spring 2017, and students in Cohort 2 were enrolled in developmental education courses in Fall 2017. This means the sample had two academic years at most to earn a credential. However, because the majority of students in the sample were not first-time college students and roughly 13% of the sample earned a credential by Spring 2019, we include this outcome in the analysis.

The primary independent variable is a dichotomous indicator variable of whether students were enrolled in a course section assigned to the treatment group, regardless of their use of the mobile apps. This phase of evaluation can therefore be described as an "intent-to-treat" analysis (Gupta, 2011). While this approach may provide a conservative estimate of the effect of using the mobile apps on student outcomes, it is necessary to avoid the threat of selection bias, namely, the likelihood that students who elected to use the apps differ systematically from students who chose not to. In addition, including all assigned students in the treatment group provides a more naturalistic estimate of the effectiveness of the treatment, given that nonuse of a supplemental instructional resource such as the mobile apps is likely in real-world contexts.

The statistical models include student-level covariates (race/ethnicity, gender, Pell receipt, age, credits attempted prior to the semester of the intervention, and credits earned prior to the semester of the intervention) and a Level 2 random intercept to account for variation in the effects of classrooms/instructors on student outcomes. Although most BPCC students had some score on a standardized assessment that was used to place them into developmental education, no more than 40% of the sample took the same assessment, and roughly 10% of the sample did not have a score for any standardized assessment. For this reason, credits attempted and earned prior to the intervention semester were used to explain variation in the outcomes and improve statistical power. Fixed effects for the 31 college by course number by instructor blocks were added to the model to account for the stratified cluster randomized design. Due to the randomized design, controlling for covariates does not substantively alter the point estimates of the treatment effect, but it does increase the power to detect statistically significant effects of the intervention.

*Statistical Models—ITT Analysis*

As the use of a cluster randomized design has the potential to result in biased estimates of the standard error of the treatment effect stemming from the introduction of Level 2 clustering (Hedges & Hedberg, 2007), multilevel modeling techniques were used to account for this clustering (Raudenbush & Bryk, 2002) with students nested in course sections. Multilevel linear regression models were used for both continuous and dichotomous outcomes. For dichotomous outcomes, linear regression was used rather than logistic regression both to facilitate interpretation and because the outcomes under study have moderate probability ranges (20%–80%) that make linear regression an appropriate choice (Hellevik, 2009). The use of logistic regression did not change the substantive findings of the analyses, and the results from logistic regression models are available upon request to the corresponding author.

Two statistical models were used. The first included fixed block effects and a fixed treatment effect to estimate the average treatment effect across course sections. The statistical equation for the model may be described as follows:

Level 1: Student Level

$$Y_{ij} = \beta_{0j} + \sum_{m=1}^{M} \beta_{1.m} X_{mij} + \varepsilon_{ij}$$

Level 2: Cluster (Course Section) Level

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(T_j\right) + \sum_{p=1}^{P-1} \gamma_{02.p} Block_{pj} + \mu_{0j}$$

$$\beta_{1.mj} = \gamma_{1.m0}$$

where $Y_{ij}$ = the *outcome* for the *i*th student in the *j*th course section; $\beta_{0j}$ = the intercept for course section *j*; $\beta_{1.mj}$ = the effects of student covariates in course section *j*; $X_{mij}$ = the *m*th of additional covariates for student *i* in course section *j*; $\varepsilon_{ij}$ = a residual error term for student *i* in course section *j*; $\gamma_{00}$ = the mean intercept; $\gamma_{01}$ = the treatment effect; $T_j$ = 1 if course section *j* is assigned to treatment, and = 0 if course section *j* is assigned to control; $Block_{pj}$ = the block number the student was enrolled in; $\gamma_{02.p}$ = the effect of block *p*; $\mu_{o,j}$ = random intercept term − deviation of course section *j*'s mean from the grand mean, conditional on covariates, assumed to be normally distributed with mean 0 and variance $\tau_{00}^2$; and $\gamma_{1.m0}$ = mean effect of student covariate *m*.

The second model estimated the extent to which the effect of treatment varies across course numbers by adding block by treatment interaction terms. The equation for this model is found below, with definitions of all terms not defined by the previous model:

*Model 2*

Level 1: Student Level

$$Y_{ij} = \beta_{0j} + \sum_{m=1}^{M} \beta_{2.m} X_{mij} + \varepsilon_{ij}$$

Level 2: Cluster (Course Section) Level

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(T_j\right) + \gamma_{02}\left(C_j\right) + \sum_{p=1}^{P-1} \gamma_{03.p} Block_{pj} + \sum_{p}^{P-1} \gamma_{04.p}\left(T_j \times C_j\right) + \mu_{0j}$$

$$\beta_{2.m} = \gamma_{2.m0}$$

where $Block_{pj}$ = 1 if the course section *j* was assigned to the treatment or comparison condition within the (randomization or matching) block *p*, and = 0 otherwise; $\gamma_{02.p}$ = the mean difference for students in course $C_j$ compared with the reference group; $C_j$ = dummy variables representing the developmental education course number (Math 98, Math 99, or English 99); $\gamma_{03.p}$ = the effect of block *p*; and $\gamma_{04.p}$ = the difference in the treatment effect for course $C_j$ and the treatment effect for the reference course, or the interaction effect of treatment with course.

*Statistical Models—TOT Analysis*

To recover the treatment effect for compliers, we use an instrumental variable (IV) approach in which we instrument for app usage based on whether a student was randomly assigned to the treatment group (Angrist et al., 1996; Angrist & Pischke, 2009). In the case of the IV approach, the treatment group includes students who were induced to use the apps after receiving login information as a result of being selected for the treatment group, whereas the counterfactual is based on similar developmental students who did not have access to the apps. Assuming the necessary conditions are satisfied, the IV approach yields an estimate of the TOT, also known as the local average treatment effect (LATE) or complier average causal effect (CACE).

We use a two-stage least squares (2SLS) approach for this analysis. In the first stage, mobile app usage is regressed on the binary treatment indicator. In the second stage, the treatment indicator in the original ITT analysis is substituted with the predicted value of mobile app usage produced by the first stage. All other covariates included in the ITT analysis are also included in the second stage of the IV model. The equations for the model may be stated as follows:

$$\left(\widehat{App}_{ij}\right) = \beta_{0j} + \gamma_{01}\left(T_j\right) \tag{1}$$

$$\left(Y_{ij}\right) = \beta_{0j} + \beta_1 \widehat{App}_{ij} + \sum_{m=1}^{M} \beta_{2.m} X_{mij} + \varepsilon_{ij} \tag{2}$$

A critical condition for a valid IV analysis is that the exclusion restriction must hold for the instrument. In this study, this means that the only way being assigned to the treatment group could affect students' college outcomes is through students actually using the mobile apps. This assumption could be violated in instances where students' outcomes are improved because their peers were given access to the apps and peer effects influence outcomes. However, because the randomization occurred at the course section level, meaning students within the same class all received access to the apps (or not) and also because only students in the treatment condition used the mobile apps, we believe this risk is minimized. Nevertheless, we consider these analyses to be more exploratory in nature compared with the confirmatory ITT analyses discussed above.

### *Effect Sizes*

We also report the estimates as effect sizes in the tables of results. Specifically, Hedges's *g* is used to compute effect sizes for continuous variables and the Cox index is used to calculate effect sizes for dichotomous variables (What Works Clearinghouse, 2020). Details on effect size calculations are included in supplemental appendix in the online version of the journal.

### *Implementation Evaluation*

The evaluation of BPCC's FITW grant also included an implementation evaluation which was categorized into two parts: an implementation fidelity evaluation and an implementation context evaluation. The implementation fidelity evaluation focused on whether BPCC had implemented the strategies outlined in its grant proposal with fidelity, whereas the implementation context evaluation focused more broadly on the factors at BPCC that influenced the degree of implementation.

The primary methods used in the implementation evaluation were annual site visits, focus groups, interviews, surveys, and regular phone calls with project staff at BPCC. The site visits occurred in both of the semesters in which the mobile apps were piloted (Spring 2017 and Fall 2017), as well as the first year of the grant before the intervention had begun and the final year of the grant when there was no active intervention (but student outcomes were still being tracked).

Approximately 40 administrators, developmental education instructors, and staff (e.g., college advisors) participated in interviews or focus groups, many of whom participated in multiple interviews or focus groups throughout the 4-year grant. Approximately 15 students participated in focus groups or interviews, but it was challenging to obtain student participation in the research. In the final year of the grant, US$20 gift cards were offered to students to participate in in-person interviews if they had used the mobile apps in their developmental education courses. However, only two students participated in these interviews that year. The response rate to the student and instructor surveys was also unacceptably low to warrant drawing conclusions from the survey data. The implementation evaluation therefore relied

TABLE 2

*Intent-to-Treat Analyses of the Effect of Mobile App Assignment on College Outcomes of Developmental Education Students*

| Estimates | Developmental education grade | Developmental education passing | Next semester persistence | Credit earned next semester | Any credential | Associate's | Certificate |
|---|---|---|---|---|---|---|---|
| Treatment | 0.136* | 0.0555** | 0.0253 | 0.458† | 0.0082 | 0.0084 | 0.0033 |
| | (0.0632) | (0.0212) | (0.0219) | (0.254) | (0.0164) | (0.0126) | (0.0159) |
| Control mean | 1.668 | 0.513 | .670 | 5.153 | 0.150 | 0.078 | 0.136 |
| Percent change | 8.1 | 10.8 | 3.7 | 8.9 | 5.5 | 10.7 | 2.4 |
| Effect size | 0.092 | 0.131 | 0.093 | 0.085 | 0.026 | 0.042 | 0.003 |
| *n* | 2,005 | 2,116 | 1,714 | 1,714 | 1,714 | 1,714 | 1,714 |

*Note.* All models control for full range of covariates and blocks. Full models are included in the supplemental appendix in the online version of the journal.
†$p < .10$. *$p < .05$. **$p < .01$.

primarily on the interviews and focus groups to gauge the factors that influenced implementation.

Finally, data from the mobile apps themselves was used to gauge implementation of the resource. This data included elements such as whether students given access to the mobile apps actually logged into the resource, the percentage of quizzes they had taken, their grades on those quizzes, and the number of videos they had watched. For brevity, the full results of the implementation evaluation will not be included, but relevant findings will be discussed below. The full results from the impact and implementation evaluation may be found in the final BPCC FITW evaluation report, which may be requested from the corresponding author.

## Results

### *ITT Analyses*

The full results of the statistical models are included in supplemental Tables A1 to A14 in the online version of the journal. For brevity, the results of the ITT analyses that estimate the ATE for the full sample are summarized in Table 2, whereas the models that examine the treatment-by-course interactions are included in Table 3. Both tables present the point estimates, standard errors, control group means, percent change in the outcome, effect sizes, and sample sizes for each of the seven models.

The largest effects are found for the outcomes specific to the developmental education courses students enrolled in. The effect size for passing the developmental education course was 0.131, with treatment students estimated to be 5.5% more likely to pass their developmental education course compared with students in the control group ($p = .009$). This effect varied across courses. Students in Math 99 were estimated to be 12.2% more likely to pass their developmental education course ($p = .001$), whereas the estimates for English 99 and Math 98 were 1.9% and 2.8%, respectively (see Table 3).

Interestingly, while the effect size for the grade students received in their developmental education courses was 0.092 and the estimated effect was statistically significant ($p = .032$), the raw difference in average grades between the treatment and control groups was not as practically significant. The average treatment effect on developmental education grades expressed as grade points was 0.136, meaning students in the treatment group earned a grade about 1.4 points higher than their control group peers on average.

The effects are more modest for the next semester outcomes. The effect sizes for the next semester outcomes of persistence (0.093) and credits earned (0.085) are roughly of the same magnitude as the effect on the grade students received in their developmental education course. The point estimate for the effect on next semester credits earned of 0.458 was marginally significant ($p = .072$) and corresponds to an 8.9% increase in credits earned given the control group mean of 5.2 credits earned the following semester. However, the estimated effect on next semester persistence of 2.5% was not significant ($p = .247$). The results provide suggestive evidence that being given access to the mobile apps

TABLE 3

*Intent-to-Treat Analyses of the Effect of Mobile App Assignment on College Outcomes of Developmental Education Students, Course Interaction*

| Estimates | Developmental education grade | Developmental education passing | Next semester persistence | Credit earned next semester | Any credential | Associate's | Certificate |
|---|---|---|---|---|---|---|---|
| Treatment (Math 99) | 0.206$^{\dagger}$ (0.111) | 0.122** (0.0373) | −0.0223 (0.0347) | 0.539 (0.408) | 0.0214 (0.0251) | 0.0137 (0.0190) | 0.0130 (0.0240) |
| English 98 | 0.758** (0.258) | 0.215* (0.0873) | −0.195* (0.0813) | −1.783$^{\dagger}$ (0.955) | 0.0138 (0.0587) | −0.0317 (0.0445) | −0.00964 (0.0563) |
| Math 98 | 0.213 (0.190) | 0.0641 (0.0614) | −0.0958$^{\dagger}$ (0.0571) | −1.189$^{\dagger}$ (0.671) | −0.0733$^{\dagger}$ (0.0413) | −0.0696* (0.0313) | −0.0580 (0.0396) |
| Treatment by English 98 | −0.189 (0.162) | −0.103$^{\dagger}$ (0.0540) | 0.0477 (0.0503) | −0.634 (0.591) | −0.0195 (0.0363) | −0.0247 (0.0275) | 0.00704 (0.0348) |
| Treatment by Math 98 | −0.0401 (0.152) | −0.0944$^{\dagger}$ (0.0504) | 0.0362 (0.0469) | −0.181 (0.551) | −0.0240 (0.0339) | −0.00178 (0.0257) | −0.0233 (0.0325) |
| *n* | 2,005 | 2,116 | 1,714 | 1,714 | 1,714 | 1,714 | 1,714 |

*Note.* All models control for full range of covariates and blocks. Full models are included in the supplemental appendix in the online version of the journal.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

may provide benefit to students in the following semester in college.

None of the estimates of the effects of the mobile apps on attainment approximated statistical significance, and all of the effect size estimates were less than 0.05. The apps were estimated to increase the attainment of any credential and associate's degrees specifically both by 0.8%, whereas the estimate for certificate attainment was only 0.3%. Once again, given the relatively short window for the evaluation, it is unknown whether larger effects would appear in the future. For example, while the estimated effect on associate's degree attainment is only 0.8%, only 7.8% of the sample earned an associate's degree by Spring 2019. This means that the percentage change in associate's degree attainment was 10.7%, which is nearly identical to the 10.8% increase in passing the developmental education course caused by assignment to using the apps. Nevertheless, the mobile apps had no discernible effect on attainment within the roughly 2-year window of the evaluation.

### *TOT Analysis*

The previous analyses estimated the ATE by comparing the outcomes of the full sample of randomized students, regardless of whether students in the treatment group actually used the apps. The following analysis use the IV approach to estimate the effect of using the mobile apps for students who are likely to use them, known as the TOT estimate or CACE. The results of the IV models for all seven outcomes are included in Table 4.

The IV analysis results in larger point estimates for every outcome analyzed, but the overall picture of the effect of the mobile apps on students' college outcomes is not significantly altered by the analysis given the lack of statistical significance for the long-term outcomes. Specifically, the CACE for students' likelihood of passing the developmental education course was 25.7 percentage points and the estimate for students' grade in the developmental education course was roughly 6 grade points, both statistically significant differences. The IV analysis also results in estimates on persistence and attainment that are all between 2.6 and 3.0 percentage points. However, none of the estimates on next semester or attainment outcomes are significant. These analyses therefore suggest that students expected to comply with the intervention by actually using the mobile apps are likely to receive significant benefits in terms of their performance in the specific developmental education course in which they are enrolled, but the effects of using the apps on long-term outcomes remain unclear.

TABLE 4

*Instrumental Variables/Treatment-on-the-Treated Analyses of the Effect of Mobile App Assignment on College Outcomes of Developmental Education Students*

| Estimates | Developmental education grade | Developmental education passing | Next semester persistence | Credit earned next semester | Any credential | Associate's | Certificate |
|---|---|---|---|---|---|---|---|
| Mobile app use | 0.596* | 0.257** | 0.0261 | 1.315 | 0.0295 | 0.0268 | 0.0277 |
| (instrumented) | (0.279) | (0.0981) | (0.0905) | (1.067) | (0.0654) | (0.0496) | (0.0627) |
| *n* | 2,005 | 2,116 | 1,714 | 1,714 | 1,714 | 1,714 | 1,714 |

*Note.* All models control for full range of covariates and blocks. Full models are included in the supplemental appendix in the online version of the journal.
$^{*}p < .05$. $^{**}p < .01$.

### *Implementation Evaluation Results*

Three findings emerged from the implementation evaluation that provide important context to the quantitative results. First, only about 20% to 25% of students in both intervention cohorts actually used the mobile apps. One cause of low engagement among students was low engagement among instructors. In interviews with instructors, a substantial proportion indicated that they had not encouraged students to use the apps in their classroom, with the biggest obstacle being the fact that they had to use different technologies and pedagogical approaches in their different classes due to the randomized research design. In addition, a number of instructors mentioned that they did not have the time or capacity to learn a new technology, particularly given that instructors at the college were required to advise students on top of their standard teaching load.

As discussed in the "Method" section, data on prior academic achievement were not consistently and uniformly collected on the sample, preventing a quantitative analysis of the relationship between students' prior achievement and their engagement with the mobile apps. However, qualitative findings suggested that there were two groups of students unlikely to use the mobile apps: high-achieving students and low-achieving or disengaged students. In contrast, students who were engaged with the course and struggling with the content, but felt capable and motivated to learn the material were more likely to use the resource. The IV analyses discussed above suggest that the apps may be quite effective for the population of students who are likely to use them.

Second, although we have generally described the intervention as a mobile app throughout, the format of the intervention evaluated in this study would be more appropriately described as a mobile-responsive website that students were instructed to save as an icon on their mobile device. Although the resource functioned like a native mobile app once it was installed in this manner, both students and instructors mentioned that the format of the intervention was confusing. In addition, instructors reported a few technical glitches in the mobile apps that could have discouraged student use, such as quizzes not submitting properly and some of the answer choices in quizzes being repeated. Although this may be common in the rollout of any new technology, it nevertheless served as an obstacle to implementation.

Third, despite the relatively modest implementation overall and the technical issues that hindered implementation, students who used the mobile apps did appear to be using them in ways aligned with their intention and the science behind the intervention. The conversation below between a student and one of the authors exemplified how students approached the apps:

Author: How have you used the apps so far?
Student: I go on there, I look at the video and I'm like, "Oh, okay. That makes sense." And then some of the questions I was like, "I don't understand this."
Author: For the quizzes you're saying?
Student: Yeah.
Author: You like the videos, but the quiz is still . . . you weren't totally sure?
Student: Yeah, it took me forever to finally get a good grade. Retry, retry, retry.
Author: Did that process help you figure out which ones were wrong or not? Like taking the quizzes multiple times?

Student: Yeah.

Author: Do you feel like that helped you retain the information or were you just like, "I'm going to just keep pressing options until I get the grade," but you're not really retaining the information?

Student: I was working it out. I would look at the problem like, "Which one did I miss." Sometimes they wouldn't even show up on the same quiz. I was like, "Okay, I see how I do this."

As highlighted in this conversation, at least some students who used the mobile apps were watching the videos, learning the content, testing their knowledge in the practice quizzes, returning to the videos to determine why they had gotten incorrect answers, and then retaking the quizzes multiple times until they received the grade they had hoped for. They also believed this process was helping them learn and retain the information. This engagement with the mobile apps was despite the fact that the actual grade students received in the quiz did not factor into their final grade. Although the scope of students' use of these practices is difficult to determine given the low responses to the student survey, data from the mobile apps suggest that students were often watching multiple videos and taking (and retaking) quizzes in a manner aligned with the principles underlying the intervention.

## Discussion

The literature is clear that students required to complete non-credit-bearing developmental education coursework are extremely unlikely to persist through college and attain a degree, particularly if they fail to complete their developmental education courses. Only 26% of community college students who enroll in developmental education courses and pass all of the courses they attempt earn an associate's degree or certificate within 6 years, and their completion rate drops to 12% if they do not pass their developmental education courses (X. Chen & Simone, 2016). At BPCC, only 14% of students enrolled in a developmental education course during Fall 2012 earned a credential by Spring 2016 (Giani, 2016).

Given the discouraging outcomes of students who begin in developmental education, a number of reforms have been devised to improve student outcomes. Reforms such as co-requisite remediation (Logue et al., 2019) and supplemental services and supports (Scrivener et al., 2015) have been found to make substantial improvements in the academic performance, persistence, and attainment of community college students. Yet these reforms often require significant investments from colleges, raising questions about their sustainability and scalability (Bettinger et al., 2013).

The use of mobile learning or "m-learning" in higher education is growing (Alrasheedi et al., 2015; B. Chen & deNoyelles, 2013; Crompton & Burke, 2018; Kaliisa & Picard, 2017; Pimmer et al., 2016). In addition to being a relatively sustainable and scalable strategy, m-learning can enable greater use of sound principles from the learning sciences in the design and delivery of content. These features include providing students with immediate and diagnostic feedback on their level of preparation, facilitating retrieval practice through ungraded quizzes, and enabling the "gamification" of learning through modularization, performance goals, and measures of progress such as "leveling up." And while online learning generally can confer these benefits, m-learning may be a particularly important strategy given the near ubiquity of mobile device ownership and the fact that students from groups historically underrepresented in higher education are more likely to own smartphones than desktop or laptop computers. Yet limited rigorous evidence exists on the impact of this strategy on students' college outcomes (Crompton & Burke, 2018). To the authors' knowledge, no studies have investigated the use of m-learning in developmental education courses using a randomized controlled trial design.

This study therefore provides the first causal evidence of the effect of providing developmental education students access to mobile apps on their academic performance, persistence, and attainment. Using a randomized controlled trial design to estimate their effectiveness, the evaluation found that students in course sections who received access to the mobile apps were significantly more likely to pass their developmental education courses and received significantly better grades compared with students who were not given access to the apps. The 5.5-percentage-point increase in developmental education passing corresponds to a 10.8% increase in the proportion of students who pass the course compared with the baseline passing rate found in the control group.

In terms of effect sizes, the treatment effect was 0.131 for passing the developmental education course and 0.092 for developmental education grades received. The size of these effects is promising, particularly for a relatively low-cost, scalable intervention (Kraft, 2020). The majority of educational interventions evaluated through a randomized controlled trial design find smaller, and often null, effects (Boulay et al., 2018; Lortie-Forgues & Inglis, 2019). Although reforms such as co-requisite remediation and the CUNY ASAP model have been found to improve students' persistence and attainment by a far greater extent, they are also often quite costly (Belfied et al., 2016; Scrivener et al., 2015). For example, the cost of the ASAP program was estimated to be US$16,284 per program group member over the 3 years of the ASAP evaluation period, or US$5,428 per year (Scrivener et al., 2015), although costs have since been reduced to roughly US$3,300 per student per year. In contrast, BPCC estimates recurring costs of roughly US$75,000 per year for maintaining the mobile apps, which now host a much larger array of content. If 2,500 students use the apps per year, the cost is roughly US$30 per student.

The effect sizes of 0.093 and 0.083 for next semester persistence and credits earned the following semester also fall in the moderate range (Kraft, 2020), although these estimates were either marginal or nonsignificant. The estimates of assignment to mobile apps on attainment were all less than 1.0% in terms of raw point estimates and less than 0.05 in terms of effect sizes. Although these effects could grow in magnitude and become statistically significant as the sample progresses through postsecondary education, the results provide no compelling evidence on the effects of the mobile apps on postsecondary attainment in the roughly 2-year analytic window used in the study.

Importantly, while overall use of the mobile apps was low, qualitative findings suggest that students who did use the apps were using them in ways congruent with self-regulated learning theory and the principles underlying the intervention. Specifically, students appeared to be engaging in the apps in ways to truly promote their learning and retention of the information and were motivated to do so despite the grades they received on the mobile quizzes not factoring into their final grades. The results of the IV analysis align with this qualitative finding. Students likely to comply with the intervention by actually using the apps were estimated to be 25.7 percentage points more likely to pass and received grades roughly 6 points higher in their developmental education courses, both statistically and practically significant differences.

Yet it is equally important to note that this intervention is not a panacea. Students in developmental education courses, and community colleges more generally, face a range of academic, financial, social, and psychological challenges that may hinder their academic performance and progress. It is imperative that community colleges and policymakers continue to address the structural, social, and financial challenges that affect student performance. Nevertheless, the use of well-designed m-learning interventions may be a sustainable and scalable strategy that community colleges can use to support the academic performance of their developmental education students.


## Acknowledgments

The authors would like to thank the faculty, staff, and students of Bossier Parish Community College for their participation in and support of the research and Cristofer Price of Abt Associates for his extremely valuable technical assistance. Questions or comments about the study can be addressed to Matt S. Giani at matt.giani@austin.utexas.edu.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the U.S. Department of Education, Fund for the Improvement of Postsecondary Education's First in the World grant program (Grant ID#: P116F150084). The results and conclusions are solely those of the authors and do not represent the views or official positions of the U.S. Department of Education.

## ORCID iD

Matt S. Giani https://orcid.org/0000-0001-7257-7253

## References

Alrasheedi, M., Capretz, L. F., & Raza, A. (2015). A systematic review of the critical factors for success of mobile learning in higher education (university students' perspective). *Journal of Educational Computing Research*, *52*(2), 257–276.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bahr, P. R. (2008). Does mathematics remediation work? A comparative analysis of academic attainment among community college students. *Research in Higher Education*, *49*(5), 420–450.

Bailey, T., Bashford, J., Boatman, A., Squires, J., Weiss, M., Doyle, W., Valentine, J. C., LaSota, R., Polanin, J. R., Spinney, E., Wilson, W., Yeide, M., & Young, S. H. (2017). *Strategies for postsecondary students in developmental education—A practice guide for college and university administrators, advisors, and faculty*. Institute of Education Sciences, What Works Clearinghouse.

Bailey, T., & Jaggars, S. S. (2016). *When college students start behind*. The Century Foundation. https://academiccommons.columbia.edu/doi/10.7916/D82B8ZHF

Bailey, T., Jeong, D. W., & Choo, S. (2008). *Referral, enrollment, and completion in developmental education sequences in community colleges*. Community College Research Center, Teachers College, Columbia University.

Barnett, E. A., Bergman, P., Kopko, E., Reddy, V., Belfield, C. R., & Roy, S. (2018). *Multiple measures placement using data analytics: An implementation and early impacts report*. Center for the Analysis of Postsecondary Readiness, Community College Research Center, Teachers College, Columbia University. https://ccrc.tc.columbia.edu/media/k2/attachments/multiple-measures-placement-using-data-analytics.pdf

Belfied, C., Jenkins, D., & Lahr, H. (2016). *Is corequisite remediation cost-effective? Early findings from Tennessee* (CCRC Research Brief No. 62). Community College Research Center, Teachers College, Columbia University. https://academiccommons.columbia.edu/doi/10.7916/D8HX1CNF

Bersin, J. (2004). *The blended learning book: Best practices, proven methodologies, and lessons learned*. John Wiley.

Bettinger, E. P., Boatman, A., & Long, B. T. (2013). Student supports: Developmental education and other academic programs. *The Future of Children*, *23*(1), 93–115.

Bloom, H. S., Bos, J. M., & Lee, S. (1999). Using cluster random assignment to measure program impacts: Statistical implications for the evaluation of education programs. *Evaluation Review*, *23*(4), 445–469.

Boulay, B., Goodson, B., Olsen, R., McCormick, R., Darrow, C., Frye, M., Gan, K., Harvill, E., & Sarna, M. (2018). *The investing in innovation fund: Summary of 67 evaluations* (Final report. NCEE 2018-4013). National Center for Education Evaluation and Regional Assistance.

Boykin, D., & Prince, A. (2015). *Inside ASAP: A resource guide on program structure, components, and management*. City University of New York.

Center for Community College Student Engagement. (2013). *A matter of degrees: Engaging practices, engaging students (high-impact practices for community college student engagement)*. Community College Leadership Program, The University of Texas at Austin.

Chen, B., & deNoyelles, A. (2013, October). Exploring students' mobile learning practices in higher education. *Educause Review*. https://er.educause.edu/articles/2013/10/exploring-students-mobile-learning-practices-in-higher-education

Chen, X., Caves, L. R., Pretlow, J., Caperton, S. A., Bryan, M., & Cooney, D. (2020). *Courses taken, credits earned, and time to degree: A first look at the postsecondary transcripts of 2011–12 beginning postsecondary students* (NCES 2020-501). National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020501

Chen, X., & Simone, S. (2016). *Remedial coursetaking at U.S. public 2- and 4-year institutions: Scope, experiences, and outcomes*. National Center for Education Statistics, U.S. Department of Education. https://nces.ed.gov/pubs2016/2016405.pdf

Crompton, H., & Burke, D. (2018). The use of mobile learning in higher education: A systematic review. *Computers & Education*, *123*, 53–64.

Edgecombe, N. D., Cormier, M. S., Bickerstaff, S. E., & Barragan, M. (2013). *Strengthening developmental education reforms: Evidence on implementation efforts from the scaling innovation project* (CCRC

Working Paper No. 61). Columbia University. https://doi.org/10.7916/D8QV3JJ3

Education Commission of the States. (2021). *50-state comparison: Developmental education policies*. https://www.ecs.org/50-state-comparison-developmental-education-policies/

Fay, M. P., Barnett, E. A., & Chavarín, O. (2017). *How states are implementing transition curricula: Results from a national scan*. Community College Research Center, Teachers College, Columbia University. https://ccrc.tc.columbia.edu/media/k2/attachments/ccrc-research-brief-how-states-implementing-transition-curricula-results-national-scan.pdf

Ganga, E., Mazzariello, A., & Edgecombe, N. (2018). *Developmental education: An introduction for policymakers*. Center for the Analysis of Postsecondary Readiness, Community College Research Center, Teachers College, Columbia University. https://ccrc.tc.columbia.edu/media/k2/attachments/developmental-education-introduction-policymakers.pdf

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *Internet and Higher Education*, *7*, 95–105.

Giani, M. S. (2016). *Bossier Parrish Community College First in the World evaluation: Baseline data analysis*. Giani Consulting & Evaluation.

Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *The Internet and Higher Education*, *18*, 4–14.

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, *2*(3), 109–112. https://dx.doi.org/10.4103%2F2229-3485.83221

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60–87.

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, *43*(1), 59–74.

Hodara, M., & Jaggars, S. S. (2014). An examination of the impact of accelerating community college students' progression through developmental education. *The Journal of Higher Education*, *85*(2), 246–276. https://doi.org/10.1080/00221546.2014.11777326

Hodara, M., Jaggars, S. S., & Karp, M. J. M. (2012). *Improving developmental education assessment and placement: Lessons from community colleges across the country* (CCRC Working Paper No. 51). Community College Research Center, Teachers College, Columbia University. https://academiccommons.columbia.edu/doi/10.7916/D8SB4F49

Hughes, K. L., & Scott-Clayton, J. E. (2011). *Assessing developmental assessment in community colleges* (CCRC Working Paper No. 19). Community College Research Center, Teachers College, Columbia University. https://files.eric.ed.gov/fulltext/ED516079.pdf

Hurwitz, M., Smith, J., Niu, S., & Howell, J. (2015). The Maine question: How is 4-year college enrollment affected by mandatory college entrance exams? *Educational Evaluation and Policy Analysis*, *37*(1), 138–159. https://doi.org/10.3102/0162373714521866

International Telecommunications Union. (2016). *ICT facts and figures 2016*. http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2016.pdf

Jackson, J., & Kurlaender, M. (2016). K-12 postsecondary alignment and school accountability: Investigating high school responses to California's Early Assessment Program. *American Journal of Education*, *122*(4), 477–503. https://doi.org/10.1086/687273

Jaggars, S. S., Hodara, M., Cho, S-W., & Xu, D. (2015). Three accelerated developmental education programs: Features, student outcomes, and implications. *Community College Review*, *43*(1), 3–26. https://doi.org/10.1177/0091552114551752

Kaliisa, R., & Picard, M. (2017). A systematic review on mobile-learning in higher education: The African perspective. *The Turkish Online Journal of Educational Technology*, *16*(1), 1–13.

Kane, T. J., Boatman, A., Kozakowski, W., Bennett, C., Hitch, R., & Weisenfeld, D. (2021). Is college remediation a barrier or a boost? Evidence from the Tennessee SAILS program. *Journal of Policy Analysis and Management*. Advance online publication. https://doi.org/10.1002/pam.22306

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. John Wiley.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253.

Kulik, J. A., Cohen, P. A., & Ebeling, B. J. (1980). Effectiveness of programmed instruction in higher education: A meta-analysis of findings. *Educational Evaluation and Policy Analysis*, *2*(6), 51–64.

Kurlaender, M., Lusher, L., & Case, M. (2020). Is early start a better start? Evaluation California state university's early start remediation policy. *Journal of Policy Analysis and Management*, *39*(2), 348–375. https://doi.org/10.1002/pam.22198

Logue, A. W., Douglas, D., & Watanabe-Rose, M. (2019). Corequisite mathematics remediation: Results over time and in different contexts. *Educational Evaluation and Policy Analysis*, *41*(3), 294–315.

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous large-scale RCTs are often uninformative: Should we be concerned? *Educational Researcher*, *48*(3), 158–166.

Miller, C., Headlam, C., Manno, M., & Cullinan, D. (2020). *Increasing community college graduation rates with a proven model*. https://www.mdrc.org/sites/default/files/ASAP_OH_3yr_Impact_Report_1.pdf

Mokher, C. G., Leeds, D. M., & Harris, J. C. (2018). Adding it up: How the Florida College and Career Readiness Initiative impacted developmental education. *Educational Evaluation and Policy Analysis*, *40*(2), 219–242. https://doi.org/10.3102/0162373717744891

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. Oxford University Press Inc.

Murray, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: Review of recent methodological developments. *American Journal of Public Health*, *94*(3), 423–432.

Nodine, T., Dadgar, M., Venezia, A., & Bracco, K. R. (2013). *Acceleration in developmental education: Game changers series*. WestEd. https://www.wested.org/resources/acceleration-in-developmental-education-game-changers-series/

Pew Research Center. (2019). *Mobile fact sheet*. http://www.pewinternet.org/fact-sheet/mobile/

Pheatt, L. E., Trimble, M. J., & Barnett, E. (2016). *Improving the transition to college: Estimating the impact of high school transition courses on short-term college outcomes* (CCRC Working Papers No. 86). Community College Research Center, Teachers College, Columbia University. https://academiccommons.columbia.edu/doi/10.7916/D8SB45QM

Pimmer, C., Mateescu, M., & Grohbiel, U. (2016). Mobile and ubiquitous learning in higher education settings: A systematic review of empirical studies. *Computers in Human Behavior*, *63*, 490–501.

Puma, M. J., Olsen, R. B., Bell, S. H., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE.

Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

St. Amour, M. (2020, June 30). Looming budget cuts threaten proven program. *Inside Higher Ed*. https://www.insidehighered.com/news/2020/06/30/experts-worry-proposed-cuts-cuny-asap-foreshadow-trend-higher-ed

Scott-Clayton, J. E. (2012). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). Community College Research Center, Teachers College, Columbia University. https://files.eric.ed.gov/fulltext/ED529866.pdf

Scrivener, S., Weiss, M. J., Ratledge, A., Rudd, R., Sommo, C., & Fresques, H. (2015). *Doubling graduation rates: Three-year effects of CUNY's Accelerated Study in Associate Programs (ASAP) for developmental education students*. MDRC. https://www.mdrc.org/sites/default/files/doubling_graduation_rates_fr.pdf

Sedlacek, W. E. (2004). *Beyond the big test: Noncognitive assessment in higher education*. Jossey-Bass.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, *24*, 86–97.

Sommo, C., Cullinan, D., Manno, M., Blake, S., & Alonzo, E. (2018, December). *Doubling graduation rates in a new state: Two-year findings from the ASAP Ohio demonstration* [MDRC, Policy Brief]. https://ssrn.com/abstract=3353494

Sommo, C., Mayer, A., Rudd, R., & Cullinan, D. (2012). *Commencement day: Six-year effects of a freshman learning community program at Kingsborough Community College*. MDRC. https://www.mdrc.org/sites/default/files/Commencement%20Day%20FR.pdf

Thorndike, E. L. (1912). *Education: A first book*. Macmillan.

Thorne, K. (2003). *Blended learning: How to integrate online and traditional learning*. Kogan Page.

Valentine, J. C., Konstantopoulos, S., & Goldrick-Rab, S. (2017). What happens to students placed into developmental education? A meta-analysis of regression discontinuity studies. *Review of Educational Research*, *87*(4), 806–833.

Venezia, A., & Hughes, K. L. (2013). Acceleration strategies in the new developmental education landscape. *New Directions for Community Colleges*, *164*, 37–45.

Visher, M. G., Weiss, M. J., Weissman, E., Rudd, R., & Wathington, H. D. (2012). *The effects of learning communities for students in developmental education: A synthesis of findings from six community colleges*. National Center for Postsecondary Research. https://files.eric.ed.gov/fulltext/ED533825.pdf

Wachen, J., Jenkins, D., Belfield, C., & Van Noy, M. (2012). *Contextualized college transition strategies for adult basic skills students: Learning from Washington State's I-BEST program model*. Community College Research Center, Teachers College, Columbia University.

What Works Clearinghouse. (2020). *Procedures handbook, Version 4.1*. Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Procedures-Handbook-v4-1-508.pdf

Xu, D., & Jaggars, S. S. (2011). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis*, *33*(3), 360–377. https://doi.org/10.3102/0162373711413814

Young, D. B., & Ley, K. (2003). Self-regulation support offered by developmental educators. *Journal of Developmental Education*, *27*(2), 2–4, 6, 8, 10.

Zachry, E., & Schneider, E. (2011). *Unlocking the gate: What we know about improving developmental education*. MDRC. https://www.mdrc.org/sites/default/files/full_595.pdf

Zeidenberg, M., Cho, S., & Jenkins, D. (2010). *Washington State's Integrated Basic Education and Skills Training program (I-BEST): New evidence of effectiveness*. Community College Research Center, Teachers College, Columbia University. https://doi.org/10.7916/D8KD262K

## Authors

MATT S. GIANI is the director of research and data science for the Office of Strategy and Policy and an assistant professor of practice in the Department of Educational Leadership and Policy at the University of Texas at Austin. His research examines how policies, programs, and interventions shape students' transitions into and through college and career

ALLISON MARTIN is the director of Institutional Effectiveness Initiatives for Bossier Parish Community College in Bossier Parish, LA, where she directs Open Educational Resource initiatives and spearheads BPCC Open Campus™ online learning platform. She is an expert in online course design, gamification, and effective learning.

Manuscript received October 13, 2019
First revision received October 5, 2020
Second revision received April 9, 2021
Accepted April 9, 2021