Who Should Take College-Level Courses?
Impact Findings From an Evaluation of a Multiple Measures Assessment Strategy

Elisabeth A. Barnett, Elizabeth Kopko, Dan Cullinan, and Clive R. Belfield
Executive Summary

Who Should Take College-Level Courses?
Impact Findings From an Evaluation of a Multiple Measures Assessment Strategy

Elisabeth A. Barnett
Community College Research Center

Elizabeth Kopko
Community College Research Center

Dan Cullinan
MDRC

Clive R. Belfield
Queens College, City University of New York

October 2020
Acknowledgments

The authors of this report are deeply grateful to the seven SUNY colleges that courageously joined this research project and have been excellent and committed partners: Cayuga Community College, Jefferson Community College, Niagara County Community College, Onondaga Community College, Rockland Community College, Schenectady County Community College, and Westchester Community College. We also greatly value our partnership with the State University of New York System Office and especially appreciate Deborah Moeckel’s support and encouragement.

Many other people have supported this work by providing feedback on drafts of this report. James Benson, our program officer at the Institute of Education Sciences, offered extensive input and useful suggestions. Peter Bergman (CCRC) was an important resource in developing our research design. Other reviewers provided helpful insights, including Thomas Brock (CCRC), Nikki Edgecombe (CCRC), Doug Slater (CCRC), Elizabeth Ganga (CCRC), and Alex Mayer (MDRC).
Overview

While many incoming community college students and broad-access four-year college students are referred to remedial programs in math or English based solely on scores they earn on standardized placement tests, large numbers of colleges have begun to use additional measures to assess the academic preparedness of entering students. Concomitant with major reform efforts in the structure of remedial (or developmental) education coursework, this trend toward the use of multiple measures assessment is informed by two strands of research: one suggests that many students traditionally assigned to prerequisite remediation would fare better by enrolling directly in college-level courses, and the other suggests that different measures of student skills and performance, and in particular the high school grade point average (GPA), may be useful in assessing college readiness.

CAPR recently completed a random assignment study of a multiple measures placement system that uses data analytics. The aim was to learn whether this alternative system yields placement determinations that lead to better student outcomes than a system based on test scores alone. Seven community colleges in the State University of New York (SUNY) system participated in the study. The alternative placement system we evaluated uses data on prior students to weight multiple measures — including placement test scores, high school GPAs, and other measures — in predictive algorithms developed at each college that are then used to place incoming students into remedial or college-level courses. Nearly 13,000 incoming students who arrived at these colleges in the fall 2016, spring 2017, and fall 2017 terms were randomly assigned to be placed using either the status quo placement system (the business-as-usual group) or the alternative placement system (the program group). The three cohorts of students were tracked through the fall 2018 term, resulting in the collection of three to five semesters of outcomes data, depending on the cohort. We also conducted research on the implementation of the alternative placement system at each college as well as a cost and cost-effectiveness analysis.

Findings from the implementation and cost components of the study show that:

- Implementation of the multiple measures, data analytics placement system was complex but successfully achieved by all the participating colleges.
- Because alternative placement resulted in many fewer enrollments in remedial courses, the total cost of using the multiple measures system was $280 less per student than using the business-as-usual system.
- Students enrolled in 0.798 fewer credits within three terms under the alternative system, saving each student, on average, $160 in tuition/fees.

Impact findings from the evaluation of student outcomes show that:

iii
• Many program group students were placed differently than they would have been under the status quo system. In math, 16 percent of program group students were “bumped up” to a college-level course; 10 percent were “bumped down” to a remedial course. In English, 44 percent were bumped up and 7 percent were bumped down.

• In math, in comparison to business-as-usual group students, program group students had modestly higher rates of placement into, enrollment in, and completion (with grade C or higher) of a college-level math course in the first term, but the higher enrollment and completion rates faded and then disappeared in the second and third terms.

• In English, program group students had higher rates of placement into, enrollment in, and completion of a college-level English course across all semesters studied. While gains declined over time, through the third term, program groups students were still 5.3 percentage points more likely to enroll in and 2.9 percentage points more likely to complete a college-level English course (with grade C or higher).

• Program group students earned slightly more credits than business-as-usual group students in the first and second terms, but the gain became insignificant in the third term. No impacts were found on student persistence or associate degree attainment.

• All gender, Pell recipient status, and race/ethnicity subpopulations considered (with the exception of men in math) had higher rates of placement into college-level courses using the alternative system. In English, these led to program group course completion rates that, compared to their same subgroup peers, were 4.6, 4.5, 3.0, and 7.1 percentage points higher for women, Pell recipients, non-Pell recipients, and Black students over three terms.

• Program group students who were bumped up into college-level courses from what their business-as-usual placements would have been were 8–10 percentage points more likely to complete a college-level math or English course within three terms. Program group students who were bumped down into developmental courses were 8–10 percentage points less likely to complete a college-level math or English course within three terms.

This study provides evidence that the use of a multiple measures, data analytics placement system contributes to better outcomes for students, including those from all the demographic groups analyzed. Yet, the (relatively few) students who were bumped down into developmental courses through the alternative system fared worse, on average, than they would have under business-as-usual placement. This suggests that colleges should consider establishing placement procedures that allow more incoming students to enroll in college-level courses.
Executive Summary

Placement testing is a near-universal part of the enrollment experience for incoming community college students (Bailey, Jaggars, & Jenkins, 2015). Community colleges accept nearly all students for admission but then make a determination about whether or not those students are immediately ready for college-level coursework. Virtually all community colleges (and more than 90 percent of public four-year colleges) use the results of placement tests — either alone or in concert with other information — to determine whether students are underprepared (Rutschow, Cormier, Dukes, & Cruz Zamora, 2019). Students deemed underprepared are typically encouraged or required to participate in remedial coursework before beginning college-level courses in those subject areas in which they are found to need academic help.

In recent years, questions have arisen about the efficacy of standardized placement tests as well as the utility of traditional developmental coursework. College practitioners and others are concerned about whether too many students are unnecessarily required to take developmental education courses before beginning college-level work. Traditional developmental courses require students to make a substantial investment of time and money, and many students who begin college by taking developmental coursework never complete a college credential (Bailey et al., 2015). Indeed, research shows that the effects of traditional developmental courses are mixed at best (Bailey, 2009; Jaggars & Stacey, 2014).

Evidence also suggests that the use of placement tests alone is inadequate in determining which students need remediation. Studies have shown that the use of multiple measures in placement decisions, and in particular the use of high school grade point average (GPA), is associated with lower rates of misplacement and higher rates of enrolling in and succeeding in college-level courses in math and English (Belfield & Crosta, 2012; Scott-Clayton, 2012). Partly in response to these findings, substantial numbers of colleges are turning to the use of multiple measures for assessing and placing students.

In 2015, the Center for the Analysis of Postsecondary Research (CAPR) began work on a random assignment study of a multiple measures, data analytics placement system to determine whether it yields placement determinations that lead to better student outcomes than a system based on test scores alone. The alternative placement system we evaluated uses data on prior students to weight multiple measures — including placement test scores, high school GPAs, and other measures — in predictive algorithms developed at each college that are then used to place incoming students into remedial or college-level courses. Seven community colleges in the State University of New York (SUNY) system participated in the study: Cayuga Community College, Jefferson Community College, Niagara Community
College, Onondaga Community College, Rockland Community College, Schenectady Community College, and Westchester Community College. A report on early findings from this research (Barnett et al., 2018) describes the implementation and costs involved in establishing such a placement system as well as the initial effects that using it had on student outcomes. The current report shares selected implementation findings but focuses mainly on providing impact findings on students during the three semesters following initial placement, as well as findings from a cost and cost-effectiveness analysis. A longer-term follow-up report on this sample of students is planned for summer 2022.

Study Design and the Implementation of an Alternative Placement System

Our study compares the effects on student outcomes of placing students into developmental or college-level courses using either a multiple measures, data analytics placement system or a status quo system that uses just one measure — placement test scores. We are also concerned with how the alternative placement system is implemented and with its costs.

Five research questions have guided the study:

1. How is a multiple measures, data analytics placement system implemented, taking into account different college contexts? What conditions facilitate or hinder its implementation?

2. What effect does using this alternative placement system have on students’ placements?

3. With respect to academic outcomes, what are the effects of placing students into courses using the alternative system compared with traditional procedures?

4. Do effects vary across different subpopulations of students?

5. What are the costs associated with using the alternative placement system? Is it cost-effective?

To answer Question 1, we conducted two rounds of implementation site visits to each of the seven colleges in which we interviewed key personnel, including administrators, staff, and faculty. To answer Questions 2 through 4, we tracked eligible students who first began the intake process at a participating college in the fall 2016, spring 2017, or fall 2017 term through the fall 2018 term. For the analyses presented in this report, student data were collected in early 2019 from the seven colleges that participated in the study and from the SUNY central institutional research office. The data allowed researchers to observe students’
outcomes for three to five semesters following placement, depending on the cohort. To answer Question 5, we conducted a study of costs as well as a cost-effectiveness analysis that incorporates outcomes data.

In order to carry out this evaluation, an alternative placement system had to be created and implemented, and random assignment procedures had to be established. Researchers and personnel at each college collaborated in these activities. We obtained 2–3 years of historical data from each college that were then used to create algorithms that weighted different factors (placement test scores, high school GPAs, time since high school graduation, etc.) according to how well they predicted success in college-level math and English courses. Faculty at each college then created placement rules by choosing cut points on each algorithm that would be used to place program group students into remedial or college-level math and English courses.

Extensive effort went into automating the alternative placement system at each college so that it could be used with all incoming students. In addition, procedures were established to randomly place about half of the incoming students (the program group) using the new data analytics system; the other half (the business-as-usual group) were placed using each college’s existing placement system (most often using the results of ACCUPLACER tests). A total of 12,971 students entered the study in three cohorts.

Overall, implementation of the multiple measures, data analytics placement system created a significant amount of up-front work to develop new processes and procedures that, once in place, generally ran smoothly and with few problems. At the beginning of the project, colleges underwent a planning process of a year or more, in close collaboration with the research team, in order to make all of the changes required to implement the alternative placement system. Among other activities, each college did the following: (1) organized a group of people to take responsibility for developing the new system, (2) compiled a historical dataset which was sent to the research team in order to create the college’s algorithms, (3) developed or improved processes for obtaining high school transcripts for incoming students and for entering transcript information into IT systems in a useful way, (4) created procedures for uploading high school data into a data system where it could be combined with test data at the appropriate time, (5) changed IT systems to capture the placement determinations derived from the use of multiple measures, (6) created new placement reports for use by students and advisors, (7) provided training to testing staff and advisors on how to interpret the new placement determinations and communicate with students about them, and (8) conducted trial runs of the new processes to troubleshoot and avoid problems during actual implementation.

While these activities were demanding, every college was successful in overcoming barriers and developing the procedures needed to support the operation of the data analytics placement system for its students. Five colleges achieved this benchmark in time for
placement of students entering in fall 2016, while the other two colleges did so in time for new student intake in fall 2017. (A fuller account of implementation findings is provided in Barnett et al., 2018.)

Data, Analysis, and Results

Sample and Method

In this experimental study, incoming students who took a placement test were randomly assigned to be placed using either the multiple measures, data analytics system or the business-as-usual system. This assignment method creates two groups of students — program group and business-as-usual group students — who should, in expectation, be similar in all ways other than their form of placement. We present aggregated findings from all participating colleges using data from three cohorts of students who went through the placement testing process in the fall 2016, spring 2017, or fall 2017 semester.

Our final analytic sample consists of 12,971 students who took a placement test at one of the seven partner colleges, of which 11,102, or about 86 percent, enrolled in at least one course of any kind between the date of testing and fall 2018. Because some students in the sample were eligible to receive either a math or an English placement rather than both, the sample for our analysis of math outcomes is reduced to 9,693 students, and the sample for analysis of English outcomes is reduced to 10,719 students. We find that differences in student characteristics and in placement test scores between program group and business-as-usual group students are generally small and statistically insignificant, which provides reassurance that the randomized treatment procedures undertaken at the colleges were performed as intended.

Our analyses were conducted using ordinary least squares regression models in which we controlled for college fixed effects and student characteristics such as gender, race/ethnicity, age, and financial aid status, as well as proxies for college preparedness.

For both math and English, we consider the following outcomes: the rate of college-level course placement (versus remedial course placement) in the same subject area, the rate of college-level course enrollment in the same subject area, and the rate of college-level course completion with a grade of C or higher in the same subject area. Because we might expect impacts to change over time, we present impact estimates for one, two, and three semesters from testing. (In the full report, we also discuss longer-term outcomes for the first cohort of students.)
Placement Determinations of Program Group Students

Because the multiple measures, data analytics placement system uses different criteria than the business-as-usual system, it could lead to more (or fewer) students being placed into college-level math or English courses. Importantly, however, any new placement procedure does not change the placement determinations of some students. Figure ES.1 shows how the placement determinations of program students differed from what they would have been under the status quo. As expected, based on prior research, the proportion of higher (or “bumped up”) placements outweighed the proportion of lower (or “bumped down”) placements in both subject areas but particularly in English, where over half of program group students were placed differently than they would have been otherwise.

Figure ES.1
Change in Placement Among Program Group Students

Main Impact Findings

As shown in Figure ES.2, placement by the algorithm increased the rate of placement into college-level math by 6.5 percentage points. But the associated gains in college-level math enrollment and completion were small and short-lived. During the first term, compared to business-as-usual group students, program group students were 2.4 percentage points (p <
more likely to enroll in a college-level math course and 2.0 percentage points ($p < .01$) more likely to pass (with grade C or higher) a college-level math course. The positive impacts on both outcomes disappeared by the third term.

**Figure ES.2**

*College-Level Math Course Outcomes (Among Students in the Math Subsample)*

![Graph showing college-level math course outcomes](image)

<table>
<thead>
<tr>
<th>Placement</th>
<th>Term 1</th>
<th>Term 2</th>
<th>Term 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business-as-usual group</td>
<td>37%</td>
<td>27%</td>
<td>15%</td>
</tr>
<tr>
<td>Program group</td>
<td>44%</td>
<td>39%</td>
<td>46%</td>
</tr>
<tr>
<td>Enrollment</td>
<td>Term 1</td>
<td>Term 2</td>
<td>Term 3</td>
</tr>
<tr>
<td>Business-as-usual group</td>
<td>29%</td>
<td>40%</td>
<td>48%</td>
</tr>
<tr>
<td>Program group</td>
<td>23%</td>
<td>24%</td>
<td>30%</td>
</tr>
<tr>
<td>Completion</td>
<td>Term 1</td>
<td>Term 2</td>
<td>Term 3</td>
</tr>
<tr>
<td>Business-as-usual group</td>
<td>17%</td>
<td>17%</td>
<td>15%</td>
</tr>
<tr>
<td>Program group</td>
<td>29%</td>
<td>29%</td>
<td>30%</td>
</tr>
</tbody>
</table>

***$p < .01$, **$p < .05$, *$p < .10$.

In English we find larger impacts across all outcomes considered. Importantly, these positive impacts in English were sustained through the third term after testing. As shown in Figure ES.3, program group students’ rate of placement into college-level English was 33.8 percentage points higher than that of business-as-usual group students. The rates of enrollment and completion among program students were also higher. Although business-as-usual group students began to catch up with program group students over time, students assigned by the algorithm maintained a modest advantage with respect to enrolling in and passing college-level English by the end of three semesters. Compared to business-as-usual group students, program group students were 5.3 percentage points ($p < .01$) more likely to enroll in a college-level English course and 2.9 percentage points ($p < .01$) more likely to pass (with grade C or higher) a college-level English course through three terms.
In addition to subject-specific impacts, we tested for impacts on overall college-level course taking, persistence, and associate degree attainment. Compared to business-as-usual group students, program group students earned, on average, 0.35 credits more college-level credits one term after testing ($p < .01$) and 0.31 more credits within the first two terms of testing ($p < .1$), but the gain became insignificant in the third term. The small, early credit impact can largely be explained by the algorithm’s effect on college-level course-taking in English, suggesting that the benefits of alternative placement did not spill over into other subjects. We find no impact on student persistence or associate degree attainment.

**Subgroup Impact Findings**

We also conducted subgroup analyses by gender (female, male), Pell recipient status (yes, no), and race/ethnicity (Black, Hispanic, White) on our main outcomes of interest in each subject: placement into, enrollment in, and completion of a college-level course. To determine whether attainment gaps between subgroups were affected by the multiple measures placement system, we also tested the significance of interaction effects between treatment status and each subgroup.
In math, we find higher rates of college-level math placement for all subgroups considered except men when placed using the algorithm \((p < .05)\). Our results suggest that the alternative placement system reversed placement gaps between female and male students: Among students in the business-as-usual group, women were less likely than men to place into college-level math; among students in the program group, women were more likely than men to place into college-level math. We also find that White students received a larger boost into college-level math from alternative placement than did their Black and Hispanic peers; that is, among students in the program group, college-level placement gaps between White and Black students and between White and Hispanic students grew larger.

Subgroup analyses in math also show that women, non-Pell recipients, and White students in the program group were 3.5, 3.8, and 3.2 percentage points \((p < .01)\), respectively, more likely to complete a college-level math course (with grade C or higher) than their same-subgroup peers in the business-as-usual group in the term following testing, but these gains were not sustained through the second or third terms. We find no evidence that existing course completion gaps by Pell recipient status changed as a result of multiple measures placement. The male-female completion gap narrowed and the White-Black completion gap widened in the first term, but these changes were not sustained in later semesters.

In English, we find much higher rates of college-level placement (of 30 percentage points or more) among program group students versus business-as-usual group students for all subgroups considered \((p < .01)\). And we find that use of the alternative placement system reversed the difference in the rate of placement into college-level English courses for women compared to men and helped to minimize the difference for Black students compared to White students.

We also find that college-level English course completion outcomes for all subgroups were higher in the first term when placed using the algorithm \((p < .01)\). These gains faded away by the third term for men and for White and Hispanic students, but they did not disappear for students in other subgroups. Although their gains declined over time, women, Pell recipients, non-Pell recipients, and Black students in the program group were 4.6, 4.5, 3.0, and 7.1 percentage points more likely than their same-subgroup peers in the business-as-usual group to complete a college-level English course (with a grade of C or higher) three terms after testing \((p < .05\) for non-Pell recipients; \(p < .01\) for all others). We do not find any evidence that gaps in the rates of course completion between related subgroups changed under the alternative placement system.

Finally, we examined outcomes of program group students whose placement determinations changed under the alternative placement system (recall Figure ES.1 showing that the placement determinations of only 26 percent of math program students and 51 percent
of English program students changed from what their business-as-usual placements would have been). We find that bumped up students had substantially better outcomes in both math and English, and that bumped down students had substantially worse outcomes. Program group students who were bumped up into college-level courses from what their business-as-usual placements would have been were 8–10 percentage points more likely to complete a college-level math or English course within three terms. Program group students who were bumped down into developmental courses were 8–10 percentage points less likely to complete a college-level math or English course within three terms.

Our findings also indicate that the college-level pass rates of program group students bumped up into college-level courses were very similar to those of students placed under the business-as-usual system. Within three terms, the status quo pass rate (with grade C or higher) in college-level math was 63 percent; the bumped-up pass rate was 60 percent. The status quo pass rate in college-level English was 67 percent; the bumped-up pass rate was 65 percent.

**Cost and Cost-Effectiveness Analysis**

To examine costs, we followed the standard approach for the economic evaluation of social programs (Levin et al., 2017). To begin, we itemized all the resources required to implement the alternative placement system and the business-as-usual system to calculate direct costs. Next we calculated the indirect costs that arise from students taking different pathways through college. To calculate cost-effectiveness (from the societal, college, and student perspectives), we identified an appropriate measure of effectiveness for each placement system. We posited that the total number of college-level credits accumulated in math and English per student after three terms would be the most valid measure of effectiveness.

The cost estimate for the alternative placement system is relative to the cost of business-as-usual testing for placement. Relative to the status quo, there are new resource requirements for the alternative system with respect to (1) administrative set-up and the collecting of data for the placement algorithms in math and English, (2) creating the algorithms, and (3) applying the algorithms at the time of placement testing. For both systems, there are costs in (4) administering placement tests. We calculated these direct costs for six colleges (resource data was insufficient at the seventh college) using the ingredients method (Levin et al., 2018).

Across the six colleges, the total cost to fully implement the new system was $958,810 (all costs are presented in present value 2016 dollars) for 5,808 students in a single cohort. However, this amount includes the cost of administering placement tests, which is estimated to have cost $174,240 for the cohort. Therefore, the net cost of implementing the alternative system was $784,560 per cohort, or $140 per student. The cost per student varied...
by college from $70 to $360 per student. This variation is primarily driven by the number of students at each college. More enrollments lead to lower costs because the costs of creating the algorithm are mostly fixed. Once the alternative placement system became fully operational, the ongoing operating costs fell substantially, to $40 per student.

To determine indirect costs and cost-effectiveness, we use the program effects on credits attempted in both developmental and college-level math and English coursework, as well as credits earned in college-level math and English courses. Program group students enrolled in 1.053 fewer developmental education credits than business-as-usual group students — or 30 percent fewer. This represents a substantial savings for both students and colleges. But program group students also enrolled in 0.255 more college-level math and English credits. In total, students placed under the alternative system attempted 0.798 fewer credits (college-level and developmental) than students placed under the status quo.

While program group students had slightly lower credit completion rates in college-level math and English courses compared to business-as-usual group students (62.6 percent vs. 63.6 percent), they attempted more college-level courses and earned more college-level credits. After three terms, program group students earned 3.975 college-level credits, and business-as-usual group students earned 3.874 such credits. Program group students thus earned 0.101 more college-level math and English credits. (Although this gain in earned credits is not statistically significant relative to business-as-usual group students, it is relevant as part of the cost-effectiveness analysis.)

Indirect costs are the costs of providing all attempted developmental and college-level credits in math and English. On average, the cost per developmental credit attempted is approximately equal to the cost per college-level credit (developmental classes are typically smaller than college-level classes, but faculty pay per class is lower). Funding per credit is divided between public support and student tuition/fees; we calculated tuition/fees as 39 percent of total expenditures per credit.

The results for this cost-effectiveness analysis from the societal or social perspective are shown in Table ES.1. The total cost of the alternative system was $280 less per student than the status quo — students took fewer developmental education credits (saving $550) that more than offset the direct cost of the alternative placement system and the extra indirect cost of providing more attempted college-level credits (at $140 and $130 respectively). The alternative placement system is more effective, given 0.101 more college-level credits earned after three terms. The cost per earned college-level credit was $1,300 for the business-as-usual system and $1,190 for the alternative placement system.
From the student perspective, the alternative placement system is clearly more cost-effective. For students, the only cost was the tuition/fees they paid for credits attempted. As students took 0.798 fewer credits under the alternative system, they saved $160. However, because students generally do not want to take developmental education, it may be more valid to focus on their developmental education savings from the alternative system. If students took 1.053 fewer developmental education credits, they saved $210 in tuition/fees (4 percent of their total spending on college).

For colleges, the determination of cost-effectiveness depends on net revenues. Colleges must pay to implement the alternative placement system; this additional cost must then be recouped by increases in net revenues (revenues over costs) from additional coursework. Estimating these costs and revenues at each college is difficult. Nevertheless, given that the alternative placement system reduced total costs and increased credit accumulation, it is plausible to conclude that it is cost-effective from the college perspective.

**Conclusion and Implications**

Colleges continue to seek ways to give students a good start in their higher education journey. The results of this study suggest that using a multiple measures, data analytics placement system is one way to increase the opportunity entering students have to succeed in college-level coursework. Some more specific lessons from this research are:

- Single placement tests are not good measures of student readiness to undertake college-level courses. As has been shown in other research, we find
that high school GPAs, especially in combination with other measures, are a better predictor of college course success.

- Colleges would be wise to set up placement systems that allow more students into college-level courses. In this study, students who were on the margin of being college-ready were much better off if they were permitted to take college courses. This can be accomplished without negatively influencing course pass rates.

- The use of a better placement system is a positive step. However, more is needed to improve student outcomes, as the impacts that occurred in this study were modest. These can include developmental education reforms as well as college-wide approaches to improving student experiences and outcomes.

This study sheds light on an important way to smooth the road for students entering college. Rather than using standardized placement tests alone, colleges can develop and deploy a multiple measures assessment and placement system that does a better job of placing students into math and English courses at a relatively low cost. The use of such a system, in tandem with other initiatives to improve student success, can make a real contribution toward improving student success in college.
References


