# The Effects and Predictive Power of the Diagnostic Assessment and Achievement of College Skills Intervention on Academic Success Indicators

Jason M. Bryer[1], Heidi L. Andrade[2], Timothy J. Cleary[3], Angela M. Lui[1], David W. Franklin Jr.[1], and Diana Akhmedjanova[2]

## Abstract

The purpose of this study was to examine the effects and predictive power of the Diagnostic Assessment and Achievement of College Skills (DAACS) on student success. DAACS is an open-source diagnostic assessment tool designed to measure newly enrolled college students' reading, writing, mathematics, and self-regulated learning skills, and to provide individualized feedback and learning resources that students can use to become better prepared for college. A randomized control trial was performed at two online colleges ($n$ = 23,467) to test the effects of DAACS on credit acquisition and retention. The results indicate an overall null effect of treatment, but post hoc analyses reveal two important findings: 1) Students who not only received the assessment results but also accessed the feedback were significantly more likely to earn credits and be retained for a second term than students who only accessed the assessment results; 2) some students who only accessed the assessment results without reading the feedback, particularly those with low scores on the assessments, low self-efficacy, or high test anxiety, had worse outcomes than the control group. We speculate that feedback mitigates the potentially negative effects of testing on student success. In addition, an examination of the predictive power of DAACS indicated that DAACS data significantly strengthen predictions of academic outcomes.

[1] City University of New York, New York, New York.
[2] University at Albany, State University of New York, Albany, New York
[3] Rutgers, The State University of New Jersey, New Brunswick, New Jersey

**Keywords**: college readiness, diagnostic assessment, remediation, predicting student success, developmental education

### Data, Materials and/or Code Availability

Analysis scripts and supplemental materials are available at https://github.com/daacs/FIPSE-Efficacy.

### Funding

The contents of this presentation were developed under grants #P116F150077 and #R305A210269 from the U.S. Department of Education. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government.

### Author Contributions

The first three authors contributed to the study conception and design. All authors contributed to material preparation. Data collection and analysis were performed by Jason Bryer and Angela Lui. All authors contributed to the writing of the first draft of the manuscript, commented on previous versions of the manuscript, and read and approved the final manuscript.

### Acknowledgements

We would like to thank the faculty, staff, and students at Excelsior College and Western Governors University who participated in this study. We would also like to acknowledge our advisory committee members who provided valuble resources and feedback on the DAACS system. List of past and current adivors is available at https://daacs.net/about/.

**The Effects and Predictive Power of the Diagnostic Assessment and Achievement of College Skills Intervention on Academic Success Indicators**

Supporting student success in college by identifying and addressing their academic preparedness is one of the biggest challenges facing higher education (Fay et al., 2017; Mokler et al., 2019; National Center for Public Policy and Higher Education & Southern Regional Education Board, 2010). Given the limitations of traditional approaches to assessing and addressing college readiness, we designed and experimentally evaluated the Diagnostic Assessment and Achievement of College Skills (DAACS), a technological platform designed to enable the diagnosis and development of essential knowledge and skills for college students. Using efficacy data from two online universities, our findings suggest that students who engage with the DAACS feedback and resources are more successful than those who do not, and the addition of DAACS data significantly increases the accuracy of identifying at-risk students.

According to The Nation's Report Card (National Assessment of Educational Progress, 2015), 75% of all high school seniors were unprepared for post-secondary coursework in mathematics, while 63% were unprepared for coursework in reading. Further, of the 1.8 million high school students who took the ACT in 2019, almost 40% failed to meet any of the four ACT College Readiness Benchmarks, which are the minimum scores required for students to have a reasonable chance of success in first-year coursework (ACT, 2019). Anticipating significant pandemic-related learning losses, especially in mathematics and science, college and university administrators are more concerned that ever about effectively meeting the needs of academically underprepared students.

One of the most common approaches to supporting students with low college preparedness involves a system of placement tests followed by remedial coursework. The logic of this approach is that by providing remedial coursework that addresses specific academic deficiencies identified by placement tests, students can transition into credit-bearing coursework. The reality, however, is that this testing-to-remediation process is fraught with inefficiencies and ineffective practices, a situation that has prompted researchers to consider alternative solutions and pathways to serving incoming freshman populations (Mokher, et al., 2019).

In this study, we conducted a randomized control trial (RCT) to examine the effects of one such alternative—the Diagnostic Assessment and Achievement of College Skills (DAACS)—on key college outcomes for first-year college students. The DAACS is a free, online, no-stakes diagnostic assessment, feedback, and intervention system. It is designed to provide students with information about their strengths and areas in need of improvement *and* to provide actionable feedback and resources they can use to optimize success (Bryer et al., 2018).

**Limitations of Placement Tests and Remedial Coursework**

For several decades, colleges have relied on placement tests to assist in the admission process and guide efforts to provide support to incoming freshmen. In recent years, researchers have demonstrated the inadequacies of these approaches and the potential negative consequences on students. For example, placement tests often generate data that do not accurately portray students' ability levels, which has led to the unnecessary and inappropriate placement of students in remedial courses (Belfield & Crosta, 2012; Scott-Clayton, 2012). Two prominent examples of traditional approaches to placement testing are ACT's COMPASS and College Board's ACCUPLACER. These tests assess students' reading, writing, and mathematics skills and inform decisions to place low-scoring students in remedial courses. Two large-scale studies showed that as many as a third of students were incorrectly placed based on their COMPASS or ACCUPLACER scores (Belfield & Crosta, 2012; Scott-Clayton, 2012): Students were either placed in remedial courses but could have passed credit-bearing classes, or they were placed in credit-bearing classes and failed. The researchers concluded that the tests have little predictive value for college success beyond what can be gleaned from high school grades. ACT eventually discontinued use of COMPASS; the College Board continues to use ACCUPLACER.

To address problems with the use of single placement tests, many colleges now rely on multiple sources of information to assess students' college preparedness, including state graduation tests, writing assessments, high school transcript information, and years since high school graduation. There is some evidence that these efforts have decreased the number of students needing to take remedial courses, perhaps in part because this additional information taps important noncognitive factors (Ganga & Mazzariello, 2019). Placement tests do not measure some skills and attributes that research has shown are necessary for succeeding in college, such as motivation, effort regulation, and the use of self-regulated learning (SRL) strategies like metacognition (Credé & Phillips, 2011; Meijer et al, 2019; Richardson et al., 2012; Weinstein et al., 1998; Wolters & Hoops, 2015; Zimmerman et al., 2011).

Finally, traditional placement tests do not provide students with feedback or access to resources they can independently use to fill in gaps in their knowledge and skills. Instead, students must pay for remedial coursework (also known as developmental courses) that does not contribute to their progress toward a degree. Remediation is very common: A Community College Research Center report noted that 68% of community college students and 40% of students at public four-year colleges took at least one remedial course, with an annual cost to all college students nationwide of approximately $7 billion (Jaggars & Stacey, 2014). The problem with remediation is that it is typically ineffective, unnecessary for the majority of students, and associated with negative outcomes such as increased cost, time to degree, and attrition (Attewell et al., 2006; Bailey & Cho, 2010; Belfield & Crosta, 2012; Rodriguez et al., 2015; Scott-Clayton & Rodriguez, 2012). Even in combination with other college supports such as learning and writing centers, remedial coursework has not adequately prepared many students entering college, with the possible exception of the lowest-scoring students at two-year institutions (Boatman & Long, 2018). Further, the placement test-remediation system has been shown to produce inequitable outcomes for minority and first-generation students, specifically regarding the disproportionate placement of these students into remedial programs (Attewell et al., 2006).

A recent study of a natural experiment in Florida highlighted the problems with remedial education. Park-Gaghan et al. (2020) studied the effects of a state-wide reform that required colleges to make developmental (i.e., remedial) courses optional for students, modify existing remedial courses to allow students to progress to credit-bearing courses more quickly, and provide enhanced advisement and support services to students. They found that after this policy was implemented, more students, especially Black and Hispanic students, took and passed introductory college courses, strongly suggesting that many of those who would have been required to take remedial courses were able to pass the credit-bearing courses. Although remedial education has an important place for some students (Park-Gaghan et al., 2020), cost effective alternatives to the placement-test-remediation approach are needed that do not rely on expensive state or federal policy changes.

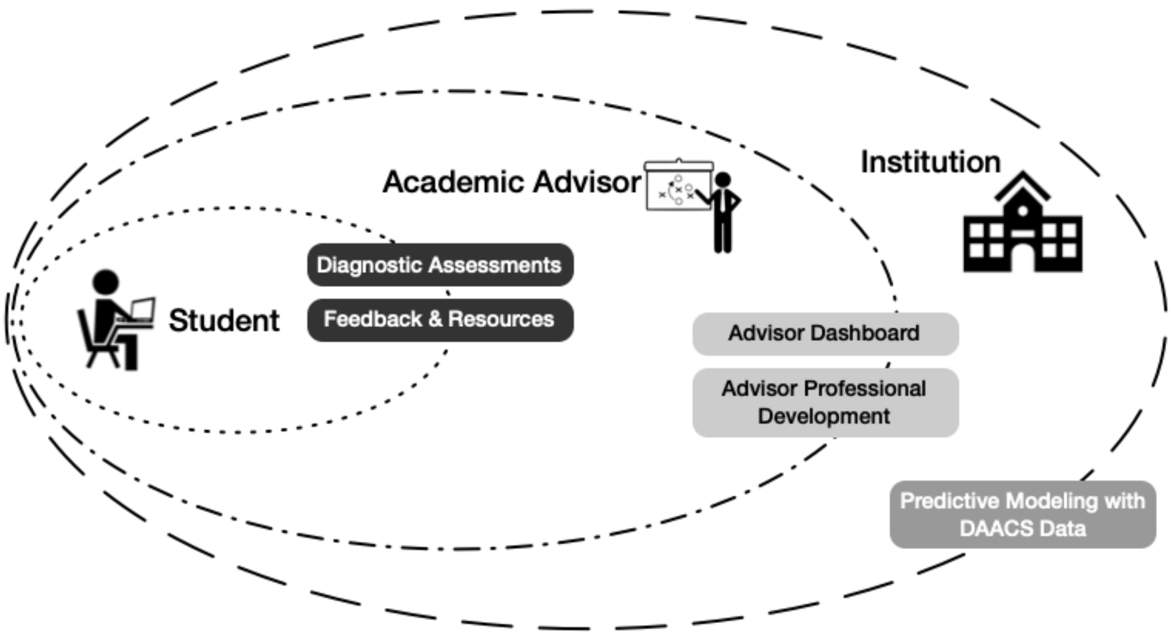
The What Works Clearinghouse (WWC) identified six empirically grounded guidelines for optimizing student preparedness and overall success in college (WWC; Bailey et al., 2016). The WWC recommends that colleges: (1) use multiple measures to assess postsecondary readiness, (2) require regular participation in enhanced advising activities, (3) compress developmental education, (4) teach students how to become self-regulated learners, (5) implement comprehensive, integrated, and long-lasting support programs, and (6) offer students performance-based monetary incentives. With the exception of monetary incentives, the DAACS system explicitly incorporates each WWC recommended practice. Through the use of open source, online, no-stakes diagnostic assessments required by the institution at which a student has enrolled, DAACS generates information about newly enrolled students' core academic, motivational, and regulatory skills. Leveraging these data, the DAACS system provides students with immediate feedback about academic strengths and areas in need of improvement and provides access to and guidance in the use of freely available resources related to building knowledge and skills on their own. The diagnostic assessment data, which is collected before students begin their coursework, can also inform advising and be used by institutions to enhance early predictions of student success or struggle.

**A New Approach: The Diagnostic Assessment and Achievement of College Skills**

DAACS has four key components: (1) diagnostic assessments, (2) individualized feedback with links to open educational resources, (3) academic advising and coaching, and (4) predictive modeling. Borrowing from Bronfenbrenner's bioecological systems theory (1979) and Bandura's social cognitive theory (2001), these components operate as reciprocal influences across distinct but related systems (see Figure 1): The student microsystem, the institutional advising mesosystem, and the institutional exosystem. DAACS was developed on the principle that any attempt to promote change in student functioning must not only be directed at the individual student, but also the contexts in which the individual learns, such as advisors, technological supports, and the interactions within and among them (Neal & Neal, 2013). DAACS also embodies social-cognitive principles, such as the importance of modeling, feedback, support systems, and student agency. Through customized feedback, access to myriad academic supports, and ongoing interactions with trained academic advisors, students are expected to develop stronger efficacy beliefs and skills in self-directing their academic lives.

**Figure 1**

*DAACS Framework and Components*




***Component 1: Diagnostic Assessments***

DAACS includes diagnostic assessments across four key areas of functioning: SRL, writing, mathematics, and reading skills. Validity and reliability evidence for each is reported in the methods section. The assessments are administered prior to students formally beginning their college coursework. Using automated scoring algorithms and procedures, students receive immediate results for each of these skill areas, presented as one, two, or three dots which roughly correspond to levels of competency—developing, emerging, and mastering, rather than numeric scores. Accompanying each competency level are individualized feedback and links to Open

Educational Resources (OERs) designed to enable students to develop knowledge and skills on their own.

**Self-regulated Learning Questionnaire.** Self-regulated learners are self-motivated, goal-directed users of effective strategies for learning and for managing competing demands on their time and attention (Zimmerman et al., 2011; Zimmerman & Schunk, 2011). The State University of New York Task Force on Remediation (2012) recommends academic supports that target students' non-cognitive skills and attitudes, including SRL and motivation beliefs, and, as noted previously, supporting the development of students' SRL is one of the six WWC recommendations. Given the importance of SRL skills in college (Efklides, 2011; Winne & Hadwin, 1998; Zimmerman, 2000), the DAACS system uses a self-report questionnaire to gather information about three key components of students' SRL: motivation, metacognition, and strategy use (Lui et al., 2018).

**Writing Assessment.** For the DAACS writing assessment, students compose an essay of at least 350 words in which they reflect on their SRL questionnaire results and commit to using strategies recommended in the feedback. Thus, the writing assessment not only assesses newly enrolled students' writing skills, but also serves as a brief intervention in support of SRL development. LightSide, an open source, automated essay scoring program (Mayfield & Rosé, 2013), was trained to score the writing assessments in terms of nine criteria related to effective college-level writing (Yagelski, 2015). Scores, feedback, and links to resources are delivered to students in one minute or less.

**Mathematics and Reading Assessments.** The mathematics and reading assessments are computer-adaptive tests with items adapted from state-mandated high school exams used to assess college readiness (Han, 2003; Jirka & Hambleton, 2005; Massachusetts Department of

Elementary and Secondary Education, 2017; New York State Education Department, 2014a, 2014b).

***Component 2: Feedback and Open Educational Resources***

Personalized feedback and OERs are two DAACS features that can help students exert agency over their own learning by getting feedback on their strengths and weaknesses and using easily accessed resources for improvement. Feedback is among the most critical influences on student achievement in general and the mastery of content and skills in particular (Andrade & Cizek, 2010; Lipnevich & Smith, 2018; Shute, 2008; Wiliam & Thompson, 2007), with an average effect size of 0.73 (Hattie, 2009). Feedback is most effective when it is detailed and narrative (not evaluative or graded), supportive, and focused on the process and self-regulation levels (Hattie & Timperley, 2007). Effective feedback provides students with concrete, actionable information about where they are (strengths and weaknesses), where they need to be (goals), and how they can move toward these goals and make improvements (next steps). Students who use DAACS are provided with immediate, individualized, actionable feedback after taking each of the four assessments.

Included with the feedback are hyperlinks to OERs that directly relate to each student's areas in need of improvement as identified by the assessments. OERs can reduce costs for higher education by making resources accessible to more students, especially to those who might not otherwise have access to them (Colvard et al., 2018; Grimaldi, et al., 2019; Zhadko & Ho, 2019). There is a modest but promising body of research on the effectiveness of OERs for improving student outcomes by increasing persistence and decreasing withdrawal rates (Grimaldi, et al., 2019; Hilton, 2016, 2020; Hilton et al., 2014). For example, Colvard and colleagues (2018) reported a decrease in failing grades and withdrawals from courses when OERs were used rather than textbooks, particularly among undergraduates from low socioeconomic backgrounds. More broadly, systematic reviews (Hilton, 2016, 2020) revealed findings that are generally in favor of OERs over traditional learning resources (e.g., textbooks) across a variety of subjects (i.e., mathematics, reading, chemistry, biology, psychology) and outcome variables (i.e., grades, test scores, credits attempted, course completion). These reviews also indicated that student and faculty perceptions of OERS are very positive, with most preferring and engaging more with OERs than traditional learning materials.

OERs are embedded into the DAACS system as a part of the actionable feedback. Two new OERs were created for DAACS: the SRL Lab (https://srl.daacs.net) and the Reading Comprehension OWL (https://owl.excelsior.edu/orc/). A library of pre-existing math related OERs was also curated. Institution-specific resources, such as the Online Writing Lab, are linked to relevant feedback as well. Sample feedback and OERs are provided in Appendix A.

***Component 3: Academic Advising***

Students in postsecondary education are typically assigned an academic advisor who assists with course planning and, through supportive dialogues, helps students troubleshoot challenges that arise (Grubb, 2001). As noted previously, a key WWC recommendation is for colleges to encourage student participation in enhanced advising activities (Bailey et al., 2016). Enhanced advising "replaces the quick, transactional structure of traditional advising (e.g., a focus on class schedules, degree requirements, and financial aid procedures) with a more holistic structure in which advisors ask deeper questions and engage more with students to help them succeed" (p. 20). With access to DAACS assessment information and knowledge gained during professional learning experiences, advisors are able to provide enhanced advising. For example, advisors had access to information about students' reading, writing, math, and self-regulated learning skills, as well as the contents of their reflective essay, in which they discussed their self-regulated learning strengths, challenges, and commitments to improve. This allowed advisors to learn more about the students and their goals and backgrounds before meeting with them, and allowed them to engage with students more meaningfully during advising sessions.

Studies that meet the WWC recommendations without reservations suggest that college students who participated in enhanced advisement were likely to accumulate more credits than students in control groups (Bailey et al., 2016; Cousert, 1999; Scrivener & Weiss, 2013; Visher et al., 2010). DAACS supports an enhanced advising experience. Through the sharing of assessment data with trained advisors via a dashboard, advisors can engage in targeted discourse with students and have a roadmap for prioritizing students' needs and connecting them to useful resources.

***Component 4: Predictive Modeling***

As institutions are increasingly being asked to do more without increasing costs, predictive models represent one innovative approach to identifying and providing supports to students who are at risk of academic failure (Baepler & Murdoch, 2010). Predictive models are a way of using statistical models with existing student information to estimate students' likelihood of success. Accurately predicting which students are least likely to succeed creates an opportunity for institutions to provide them with targeted supports and resources, thereby allocating limited resources more effectively. DAACS is designed to be integrated into student information systems to prioritize outreach to potentially at-risk students. This study included an examination of the accuracy of predictive models that include students' DAACS assessment results in addition to demographic variables and other academic performance indicators such as transfer credit and prior college experience.

**Purpose of the Study**

The purposes of this study were to examine the effectiveness of the DAACS system (i.e., information and feedback on college readiness, access to relevant OERs, and advising and predictive models) in improving early college outcomes (e.g., credit acquisition) for students at two institutions and to determine whether the data generated by the DAACS system enhanced the accuracy of predictive analytic models of academic success. The study addresses three research questions:

1. What is the overall effect of DAACS on students' early credit acquisition and academic achievement?
2. Is there a dosage effect in terms of student access to DAACS feedback and resources?
3. To what extent does the inclusion of DAACS results in predictive analytic models increase the accuracy of predictions of academic success?

Given the theoretical and empirical foundation for the various components of the DAACS system, we hypothesized that students who used DAACS upon entering college would attain better college outcomes than students who did not. We also expected a dosage effect in terms of the number of times treatment students accessed their DAACS feedback and resources. Because DAACS gathers information about skills and processes that have been shown to predict student success in college, we also expected that the inclusion of DAACS data would improve predictive models of student success.

## Method

To meet the What Works Clearinghouse (WWC) standards without reservation (Bailey et al., 2016), a randomized control trial was used. At two online institutions of higher education (Excelsior College and Western Governors University), newly enrolled students were randomly assigned to one of two orientations. Students enrolled in the study over a continuous one-year period, so randomization occurred at the time of enrollment for each student. We used the oddness or evenness of student identifier numbers to determine the orientation group into which the student would be placed: even IDs were placed into treatment, and odd IDs were in control. The control group received business-as-usual (BAU), while students in the treatment group were expected to take the DAACS as a part of their online orientation program.

**The Context**

All students at Western Governors University (WGU) were required to complete orientation before enrolling for credit and did so. Due to institutional restructuring at Excelsior College (EC) at the time of the study, however, it was not possible to require students to complete DAACS prior to orientation enrollment. Thus, there was no consequence for students not completing the online orientation.

EC and WGU differed in terms of the infrastructure for supporting and advising students. WGU has a competency-based program whereby students enroll in a six-month term where they are expected to complete at least 12 credits. Students are assigned a mentor/academic advisor with whom they meet approximately every two weeks to check on their progress. Students work independently but have access to subject matter experts as they complete assignments or "competencies". EC, in contrast, employs a more traditional course-based model where students enroll in eight or 15-week courses. Each course has an instructor and a maximum of 25 students. Courses are organized in weekly modules and coursework is completed asynchronously. At the time of this study, students at EC rarely met with their academic advisors, typically only to address administrative tasks such as course registration.

**Participants**

The sample included 23,467 incoming undergraduate students from Excelsior College and Western Governors University, two private, nonprofit, online colleges in the United States. Students enrolled between April 15, 2017, and December 31, 2017. Students included in the study were at least 18 years old and enrolled in an Associates or Baccalaureate degree program. EC ($n$ = 10,285) has an open enrollment policy; WGU ($n$ = 13,182) requires students to take an admissions test. Both institutions serve predominately non-traditional, first-generation college students with an average age of 35.22 (SD = 9.26) at EC and 33.61 (SD = 8.91) at WGU. Students were enrolled in one of six programs at EC, and one of 30 programs at WGU. The number of credits that were transferred in for students at EC ranged from 2 to 229, with a median of 48; transfer credits for students at WGU ranged from 0 to 141, with a median of 32. Additional demographic information is provided in Table 1.

**Table 1**

*Sample Demographics by Institution and Experimental Condition*

| | | Excelsior | | | | WGU | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Control (*n*=5144) | | Treat (*n*=5141) | | Control (*n*=6646) | | Treat (*n*=6536) | |
| | | *N* | % | *n* | % | *n* | % | *n* | % |
| Gender | Female | 2052 | 40 | 2116 | 41 | 3868 | 58 | 3758 | 57 |
| | Male | 3092 | 60 | 3025 | 59 | 2775 | 42 | 2773 | 42 |
| | Missing | -- | -- | -- | -- | 3 | 0 | 5 | 0 |
| Ethnicity | Am. Indian or Alaskan Native | 32 | 1 | 31 | 1 | 63 | 1 | 52 | 1 |
| | Asian | 142 | 3 | 189 | 4 | 188 | 3 | 190 | 3 |
| | Black | 1045 | 20 | 1045 | 20 | 691 | 10 | 747 | 11 |
| | Hispanic | 670 | 13 | 697 | 14 | 157 | 2 | 197 | 3 |

| | | Excelsior | | | | WGU | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Control | | Treat | | Control | | Treat | |
| | | (*n*=5144) | | (*n*=5141) | | (*n*=6646) | | (*n*=6536) | |
| | Native Hawaiian | 42 | 1 | 47 | 1 | 42 | 1 | 45 | 1 |
| | Multiple | 226 | 4 | 218 | 4 | 283 | 4 | 246 | 4 |
| | White | 2932 | 57 | 2852 | 55 | 5033 | 76 | 4907 | 75 |
| | NA | 55 | 1 | 62 | 1 | 192 | 3 | 152 | 2 |
| First-Generation | No | 1668 | 32 | 1737 | 34 | 3969 | 60 | 3889 | 60 |
| | Yes | 1766 | 34 | 1731 | 34 | 2677 | 40 | 2647 | 40 |
| | NA | 1710 | 33 | 1673 | 33 | -- | -- | -- | -- |
| Military | No | 2247 | 44 | 2338 | 45 | 5908 | 89 | 5815 | 89 |
| | Yes | 2897 | 56 | 2803 | 55 | 738 | 11 | 721 | 11 |
| Employment Status | Not-employed | 844 | 16 | 788 | 15 | 863 | 13 | 841 | 13 |
| | Part-time | 307 | 6 | 352 | 7 | 899 | 14 | 822 | 13 |
| | Full-time | 3764 | 73 | 3773 | 73 | 4507 | 68 | 4511 | 69 |
| | NA | 229 | 5 | 228 | 4 | 377 | 6 | 362 | 6 |
| English Language Native | No | 359 | 7 | 365 | 7 | -- | -- | -- | -- |
| | Yes | 4747 | 92 | 4716 | 92 | -- | -- | -- | -- |
| | NA | 38 | | 60 | | | | | |
| Citizenship Status | Non-citizen | -- | -- | -- | -- | 94 | 1 | 60 | 1 |
| | Non-resident alien | -- | -- | -- | -- | 37 | 1 | 32 | 0 |
| | U.S. citizen | -- | -- | -- | -- | 6409 | 96 | 6368 | 97 |
| | NA | | | | | 106 | 2 | 76 | 1 |
| Degree Level of Program Enrolled | Associates | 2102 | 41 | 2180 | 42 | 0 | 0 | 0 | 0 |
| | Bachelor | 3042 | 59 | 2961 | 58 | 6646 | 100 | 6536 | 100 |
| Income | < $25,000 | 429 | 8 | 448 | 9 | 1207 | 18 | 1187 | 18 |
| | < $35,000 | 518 | 10 | 433 | 8 | 869 | 13 | 906 | 14 |
| | < $45,000 | 456 | 9 | 476 | 9 | 821 | 12 | 863 | 13 |

| | Excelsior | | | | WGU | | | |
|---|---|---|---|---|---|---|---|---|
| | Control | | Treat | | Control | | Treat | |
| | (*n*=5144) | | (*n*=5141) | | (*n*=6646) | | (*n*=6536) | |
| ≥ $45,000 | 2061 | 40 | 2128 | 41 | 3325 | 50 | 3151 | 48 |
| NA | 1680 | 33 | 1656 | 32 | 424 | 6 | 429 | 7 |

Note. Percentages may not total 100 due to rounding.

**Design and Procedures**

An RCT design was used, whereby students were randomly assigned to one of two versions of an online, self-paced orientation course for newly enrolled students at each institution: treatment with DAACS (n = 11,677) or control without DAACS (n = 11,790). All incoming students at both colleges were expected to complete orientation before beginning coursework. Students in the treatment condition were expected to take all four DAACS assessments as part of orientation. However, because the orientation requirement was not enforced at EC, 2,088 (43%) students in the treatment group at EC did not complete DAACS. Students at both institutions who completed the assessments received individualized feedback and suggestions through the DAACS website for how to address gaps in their knowledge and improve their skills (see Appendix A). They were only required to complete the assessments – viewing and using the results and feedback were expected but optional.

At both EC and WGU, students in the control condition were assigned to an orientation that was identical to that of the treatment group but without the DAACS components. That is, students in the control group did not take the DAACS assessments or receive DAACS feedback, and only interacted with advisors who were not trained to use DAACS.

At WGU, students in the treatment group were also assigned to an academic advisor trained in the operation, use, and interpretation of DAACS-related information, but fewer than 36% of students had an academic advisor view their DAACS results. At EC, due to institutional restrictions at the time the study began, no academic advisors utilized DAACS results when advising students. Given that the advising component of DAACS was not fully realized in this study, we do not include data or analyses related to advising in this paper.

**Data Sources**

***Self-regulated Learning Questionnaire***

The SRL questionnaire consists of 47 Likert-type items adapted from established SRL measures (Cleary, 2006; Driscoll, 2007; Dugan & Andrade, 2011; Dweck, 2006; Schraw & Dennison, 1994). The items are organized in three domains: (1) *motivation,* which measures students' mastery orientation, test anxiety, self-efficacy, and mindset; (2) *learning strategies,* which measures students' help-seeking behaviors, as well as the ability to manage their understanding, time, and environment; and (3) *metacognition,* which measures students' ability to plan, monitor, and evaluate their learning.

The SRL questionnaire has good psychometric qualities. Factor-analytic evidence of construct validity was obtained through second-order confirmatory factor analysis (CFA, $N$ = 6644). It showed that all 11 scales (i.e., first-order factors), individually, had moderate to high internal consistency reliability estimates of Cronbach's alpha: $.69 \leq \alpha \leq .91$. The second-order internal consistencies of the three domains were also acceptable: ($.61 \leq \alpha \leq .89$). The fit of the second-order CFA model to the collected data was estimated at CFI = 0.868 and RMSEA [90% CI]: 0.053 [0.052, 0.054], suggesting that 1) the items clustered well as domains and scales, and 2) inferences drawn from the questionnaire scores are valid and reliable (Lui et al., 2018).

***Writing Assessment***

The writing assessment asks students to summarize their SRL questionnaire results, identify specific strategies for improving their SRL, and commit to using them. LightSide

(Mayfield & Rosé, 2013), an open source, automated essay scoring program, was trained to score the writing assessments in terms of nine criteria related to effective college-level writing (Summary, Suggestions, Structure, Transitions, Cohesion, Focus on the Main Idea, Grammatically Correct Sentences, Sentence Complexity, and Conventions; Yagelski, 2015). Students received feedback within one minute. The first 1,093 essays were scored by human raters, and 597 (55%) were double scored and adjudicated. The essays and their scores were used to train LightSide.

Exact percent agreement between human raters ranged from 55% to 63%; exact percent agreement between human and LightSide ratings ranged from 47% to 74% (Akhmedjanova et al., 2019). LightSide and human scorers rarely disagreed by more than one level on the three-level rubric, but because agreement did not always reach the acceptable threshold of 70%, two expert raters agreed on scores for 10 essays and compared them to LightSide scores. Average exact agreement was 61.1%, and average adjacent agreement was 96.8% (Author, 2019).

We consider the human-machine scoring agreement levels to be low but sufficient for an experimental, no-stakes diagnostic assessment. In a poll of students, 38 of 40 (95%) said the writing assessment was mostly or completely accurate. In addition, the CFA model using a human-scored sample of randomly selected students' essays ($N$ = 879) had a good fit to data ($\chi^2/df$ = 4.74, CFI = .966, and RMSEA [90% CI] = .065 [.052, .079], suggesting that the factor structure of the scores generally conform to our conceptual framework as represented by the rubric (Akhmedjanova et al., 2019; Andrade et al., 2018). Further, Yu and colleagues (2022) provide predictive validity evidence for the writing assessment, showing that students with the lowest DAACS writing scores were approximately three times less likely to pass their first writing course than students with the maximum total writing.

### *Mathematics and Reading Assessments*

The DAACS mathematics assessment includes 122 items that cover word problems, geometry, variables and equations, numbers and calculations, and lines and functions. Using a computer adaptive testing framework, students saw a stratified random selection of 18 to 24 items among the five domains, with item difficulty tailored to the individual student.

The DAACS reading assessment includes 30 passages of varying difficulty levels with six items per passage. The items assess students' understanding of the ideas, language, purpose, and structure of written passages, as well as their ability to make inferences. Students saw a random selection of three to four passages with difficulty levels tailored to their performance on the first selection.

Cronbach's alphas for the mathematics and reading assessments were 0.67 and 0.69, respectively, suggesting that the reliability of scores from these assessments is adequate. Criterion validity evidence was also gathered via inter-scale and intra-scale correlations. Appendix B presents the average inter-scale correlation between the math and reading total scores (rho = .37), as well as inter-scale correlation. Fifty-nine of 62 correlation coefficients have values between .20 and .40, which suggests that the math and reading scales are reasonably homogeneous.

### *Academic Success*

Two indicators of student academic success were used to determine the effectiveness of DAACS: (1) *Success rate* is the ratio of credits earned to credits attempted in their first term; (2) *Term-to-term retention* is a dichotomous variable that indicates whether students enrolled in courses the following term. This latter operational definition is adapted from the definition of retention rate conveyed by the Institute of Education Sciences (IES) National Center for

Education Statistics (NCES): “the percentage of first-time degree/certificate-seeking students from the previous fall who either re-enrolled or successfully completed their program by the current fall” (de Brey et al., 2022, para. 4).

***DAACS System Analytics on Dosage***

To address the first research question regarding intent-to-treat, a dichotomous treatment variable was created to indicate whether students were assigned to the treatment group (DAACS) or the business as usual control group. As noted above, approximately 40% of students in both the treatment and control groups at EC did not complete orientation and, as a result, many students in the treatment group at EC did not receive the DAACS intervention. Nonetheless, in order to adhere to the WWC standards of analyzing data based upon intent-to-treat, all students who enrolled and were assigned to either group were included in this analysis regardless of their actual DAACS usage.

To address the second research question regarding dosage, a variable called *DAACS dosage* was created to indicate whether students who took the assessments viewed the feedback pages and/or visited the resources recommended there (the OERs*)*. DAACS dosage is operationalized as the number of times a student clicked on the various links within the DAACS feedback system: (1) students who completed the assessments without viewing DAACS feedback (Assessment Results Only) and (2) students who took the assessments and viewed at least one of the feedback pages (Assessment Results + Feedback). We provide sample sizes for each of these groups in Table 2.

**Table 2**

*DAACS Dosage by Institution*

| DAACS Dosage | | EC* | WGU* |
|---|---|---|---|
| Treatment | | 3,083 | 6382 |
| | Assessment Only | 1,610 | 2,551 |
| | Assessment + Feedback | 1,473 | 3,831 |
| Control | | 3,138 | 6,646 |

**Note*: EC students who were in the treatment group but did not attend orientation and therefore did not complete DAACS were omitted ($n_T$ = 2088; $n_C$ = 2042). WGU students who did not complete DAACS were also omitted ($n$ = 154).

Students who did not access orientation at both institutions were excluded from this variable because the focus of this analysis is on DAACS dosage. In addition, for EC, there were a significant proportion of students in the treatment and control groups who did not access any part of the orientation course. The proportions of students not accessing orientation were statistically equivalent in the treatment and control groups. Moreover, there was no statistically significant difference in either the demographic or outcome variables between students in the control and treatment groups who did not access orientation. Given these results, students who did not access orientation were excluded from the analysis.

**Data Analytic Plan**

Prior to data analyses, we examined the treatment and control groups for baseline equivalence. Examination of key covariates revealed baseline inequivalence across several variables at EC (Gender, Income, English language native, Employment, Age, Military status, Program) and WGU (Citizenship status, Ethnicity, Income, Program). For this reason, we followed WWC guidelines (2017) and conducted our analyses using inverse probability weights

(IPWs). However, because the results with and without the use of IPWs were essentially equivalent, we report on the results without IPWs, which are easier to interpret.

Table 3 summarizes the analytic procedures and variables used for each research question. For Research Question 1, we examined the overall effects of DAACS using the intent-to-treat conditions (randomized treatment or control) and two outcome variables (one dichotomous and one continuous) for each institution. Chi-squared tests were used for the dichotomous outcome and independent sample *t*-tests for the continuous outcome. Unless otherwise noted, statistical significance was tested at $p < 0.05$ with Bonferroni adjustments (0.05/5 comparisons = 0.01) to account for multiple comparisons and minimize Type 1 error.

**Table 3**

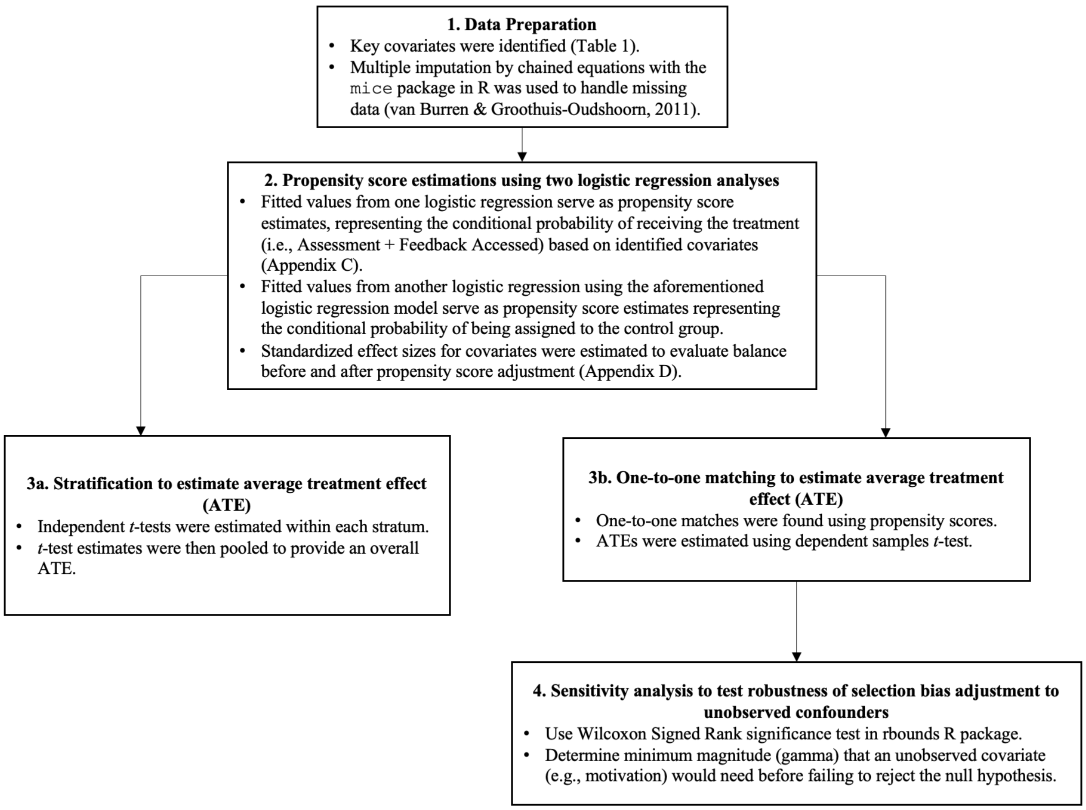
*Data Sources and Analytic Procedures by Research Questions*

| Research Question | Independent & Predictor Variables | Outcome Variables | Analytic Procedures |
|---|---|---|---|
| 1. What is the overall effect of DAACS on students' early credit acquisition and academic achievement? | ▪ Intent-to-treat conditions (i.e., Treatment/Control) | ▪ Term-to-term retention (No/Yes)<br>▪ Success rate (ratio of credit earned to credit attempted) | Null hypothesis testing |
| 2. Is there a dosage effect in terms of student access of DAACS feedback and resources? | ▪ DAACS dosage: Results Only or Results + Feedback Accessed | ▪ Term-to-term retention (No/Yes)<br>▪ Success rate (ratio of credit earned to credit attempted) | Propensity score analysis and Complier Average Causal Effect |
| 3. To what extent does the inclusion of DAACS results in predictive analytic models increase the accuracy of predictions of academic success? | ▪ Demographics<br>▪ DAACS SRL, writing, math, and reading results<br>▪ DAACS dosage | ▪ Term-to-term retention (No/Yes) | Predictive modeling |

For Research Question 2, we used propensity score analysis (PSA; Rosenbaum & Rubin, 1983) to adjust for student self-selection of DAACS dosage: Assessment Results Only compared to Assessment + Feedback. PSA is a quasi-experimental design that has been shown to provide causal estimates from observational data consistent with randomized control trials (Lonjon et al., 2014; Shadish et al., 2008; Stuart et al, 2017). In accordance with Leite (2017), we followed the steps of PSA illustrated in Figure 2.

**Figure 2**

*Steps of Propensity Score Analysis to Examine Effects of DAACS Dosage on Student Success*




After data preparation, logistic regression was used with the identified covariates to estimate propensity scores (see Appendix C); the fitted values serve as the propensity scores, which are defined as the conditional probability of receiving the treatment (defined as Assessment + Feedback). We also estimated propensity scores for students assigned to the control group using the logistic regression model estimated with students assigned to the treatment. With this design we were able to compare students who completed the assessments and viewed feedback to the students who only took the assessments, as well as to students assigned to the control group. To determine whether balance was achieved, the standardized effect size was estimated for each covariate before and after propensity score stratification (Appendix D). After propensity score adjustment, the effect sizes were all substantially reduced, with the largest effect size being 0.07, and the majority being less than 0.01.

Once propensity scores were generated, average treatment effects were estimated using two methods (i.e., stratification and matching) to compare the outcomes of Assessment Results + Feedback students to Assessment Results Only as well as to Control. The Matching package in R (Sekhon, 2011) was used to find one-to-one matches using dependent sample *t*-tests. The PSAgraphics package in R (Helmreich & Pruzek, 2009) was used to estimate average treatment effect (ATE) with independent *t*-tests within each stratum.

For the matching methods, sensitivity analysis (Rosenbaum, 2002) was conducted using the rbounds (Keele, 2022) package in R to test the robustness of the selection bias adjustment to unobserved confounders (Rosenbaum, 2005). The goal of this statistical technique is to determine the minimum magnitude (gamma) that an unobserved covariate would need before failing to reject the null hypothesis. The rbounds package uses a Wilcoxon Signed Rank significance test, which relaxes the normality assumptions. Since sensitivity analysis is only well defined for matching methods, Rosenbaum (2012) has recommended testing the null hypothesis twice. This procedure also tests the sensitivity of the results to the method of choice, and justifies our use of two methods – matching and stratification – to address Research Question 2.

To address Research Question 3 regarding predictive analysis, we used the dichotomous outcome variable, term-to-term retention (Retention). Multiple imputation by chained equations with the mice package in R was used to handle missing data (van Burren & Groothuis-Oudshoorn, 2011). At WGU, household income and employment status had the most missingness, with 6.5% and 5.7% missing, respectively. Missing data for income range and first-generation status was more serious at EC, each of which had 33% missingness. Investigation into this missingness revealed that these data were not missing at random: Some students could bypass an enrollment form that asked for these data. However, because the intent of this analysis was to compare the relative performance of model accuracy, and any bias would be present in all models, imputation was an acceptable method (Sinharay et al., 2001). The remaining demographic variables had fewer than 5% missing data.

The datasets from each institution were used to evaluate the accuracy of predictive models using logistic regression, classification trees, and random forest (Breiman, 2001). Since random forest outperformed logistic regression and classification trees in all but one instance, we only report the results from random forest. The use of random forest is supported by a study conducted by Fernández-Delgado et al. (2014) that examined 179 classifiers with 121 datasets that concluded random forest ranked the highest among the classifiers.

## Results

### Research Question 1: What is the Overall Effect of DAACS on Early Credit Acquisition and Academic Achievement?

As noted previously, this question was addressed using an intent-to-treat (ITT) approach to the analyses. All students who were randomly assigned to the treatment group were considered to be treated, although 43% of students at EC did not receive the intervention. Given this latter issue the high percentage of students not receiving DAACS, it was not surprising that neither the chi-square analyses of the effects of the intervention on term-to-term retention (Retention) nor the *t*-tests of the ratio of credits earned to credits attempted (Success Rate) provided evidence of group differences at either of the two institutions (Table 4).

**Table 4**

*Estimated Causal Effects for Intent-to-Treat: Chi-Square Analyses of Retained and t-tests of Success Rate*

| | | Excelsior College | | | | | Western Governors University | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Outcome | | Treatment ($n$ = 5141) | Control ($n$ = 5144) | test-statistic | $p$ | *Effect size^* | Treatment ($n$ = 6542) | Control ($n$ = 6658) | test-statistic | $p$ | *Effect size^* |
| Retained | Yes | 39.8% | 39.0% | $\chi^2 = 1.27$ | .26 | 0.01 | 73.9% | 74.2% | $\chi^2 = 0.17$ | .68 | 0.004 |
| | No | 60.2% | 61.0% | | | | 26.1% | 25.8% | | | |
| Success Rate | | M = .60 | M = .60 | $t_{10283} = 0.25$ | .80 | 0.005 | M = .78 | M = .78 | $t_{13179} = 0.02$ | .99 | 0.000 |

Note. OTP = On-time progress

^Cohen's $d$ is reported for Success Rate; phi coefficient for Retained

*$*p < .05$*

## Research Question 2: Is There an Effect of Dosage on Success Rate and/or Retention?

Only 54% of students at EC and 40% of students at WGU who engaged with DAACS opted to access feedback beyond the results page. This variability in students' usage of the DAACS feedback and OERs naturally created two dosage-related subgroups that could be compared to the control group: 1) Assessment Results Only, and 2) Assessment Results + Feedback (see Table 2). The correlations between the outcomes and the log of page views of feedback/OERs (our proxy for dosage) in the treatment groups range from 0.07 to 0.14 (see Table 5). All correlations are statistically significant at $p < 0.01$, suggesting there is a relationship between dosage and student success.

**Table 5**

*Correlations between Outcome Measures and Dosage (Log of Page Views)*

| | Excelsior College | | | Western Governors University | | |
|---|---|---|---|---|---|---|
| Outcome | Correlation | Test-Statistic | $p$ | Correlation | Test-Statistic | $p$ |
| Retention | 0.1 | $t_{10,249} = 10.04$ | < 0.01 | 0.07 | $t_{6,384} = 5.33$ | < 0.01 |
| Success Rate | 0.14 | $t_{10,249} = 14.08$ | < 0.01 | 0.1 | $t_{6,384} = 8.17$ | < 0.01 |

Propensity score analysis was used to examine the effect of dosage on success rate and retention. Propensity scores were estimated using logistic regression with students assigned to the treatment. Using this model, propensity scores were also estimated for students assigned to the control group (see Figure 2). For matching, one-to-one matches were found using the propensity scores and ATE were estimated using a dependent sample *t*-test. For stratification, six strata were used across propensity scores. Independent sample t-tests were performed for each strata and pooled to provide an overall ATE using the PSAgraphics R package (Helmreich & Pruzek, 2009).

Table 6 provides the ATEs for Retention and Success Rate for both institutions using matching and stratification. There is a positive effect of Assessment Results and Feedback compared to Results Only for both methods on both measures at both institutions. When comparing Assessment Results and Feedback to Control, the positive effect remains for Retention and Success Rate at EC but only for Success Rate at WGU.

**Table 6**

*Estimated Causal Effects for Propensity Score Analyses (Matching and Stratification)*

| | | Excelsior College | | | | Western Governors University | | | |
|---|---|---|---|---|---|---|---|---|---|
| Outcome | Group | ATE | *t*-statistic | *p* | Gamma | ATE | Statistic | *p* | Gamma |
| | | | | | Matching | | | | |
| Retention | Results + Feedback vs. Results Only | **0.08** | $t_{3082} = 6.59$ | < 0.01 | 1.30 | **0.04** | $t_{6381} = 4.78$ | < 0.01 | 1.10 |
| Retention | Results + Feedback vs. Control | **0.04** | $t_{4610} = 4.10$ | < 0.01 | 1.18 | 0.00 | $t_{10476} = 0.69$ | 0.48 | |
| Retention | Results Only vs. Control | **-0.03** | $t_{4747} = -2.53$ | 0.01 | 1.50 | **-0.03** | $t_{9196} = -4.14$ | < 0.01 | 1.50 |
| Success Rate | Results + Feedback vs. Results Only | **0.04** | $t_{3082} = 3.98$ | < 0.01 | 1.25 | **0.07** | $t_{6381} = 11.36$ | < 0.01 | 1.33 |
| Success Rate | Results + Feedback vs. Control | **0.03** | $t_{4610} = 2.99$ | < 0.01 | 1.19 | **0.02** | $t_{10476} = 5.06$ | < 0.01 | 1.08 |
| Success Rate | Results Only vs. Control | -0.01 | $t_{4747} = -0.96$ | 0.33 | | **-0.04** | $t_{9196} = -8.38$ | < 0.01 | 1.50 |
| | | | | | Stratification | | | | |
| Retention | Results + Feedback vs. Results Only | **0.07** | $t_{3071} = 3.93$ | < 0.01 | | **0.03** | $t_{6370} = 2.54$ | 0.01 | |
| Retention | Results + Feedback vs. Control | **0.07** | $t_{3071} = 3.93$ | < 0.01 | | 0.01 | $t_{10465} = 0.82$ | 0.41 | |
| Retention | Results Only vs. Control | -0.02 | $t_{4736} = -1.37$ | 0.17 | | -0.02 | $t_{9185} = -1.97$ | 0.05 | |
| Success Rate | Results + Feedback vs. Results Only | **0.04** | $t_{3071} = 2.61$ | 0.01 | | **0.06** | $t_{6370} = 7.15$ | < 0.01 | |
| Success Rate | Results + Feedback vs. Control | **0.03** | $t_{4599} = 2.47$ | 0.01 | | **0.02** | $t_{10465} = 3.88$ | < 0.01 | |
| Success Rate | Results Only vs. Control | -0.01 | $t_{4736} = -0.58$ | 0.56 | | **-0.04** | $t_{9185} = -4.37$ | < 0.01 | |

The results in Table 6 also indicate that the ATE for Assessment Results Only compared to Control are negative; half of them are statistically significant. This unexpected finding, although inconsistent, suggests that students who received the assessment results but chose not to view the feedback were less likely to succeed than students in the control group. Post-hoc analyses indicate that at WGU, adjusting for SRL, treatment, and all demographic variables using propensity scores, test anxiety is the only SRL variable that predicts Retention and Success Rate using propensity score weights ($\beta_{\text{Retention}} = 0.13$, $t_{6381} = 3.19$, $p < 0.01$; $\beta_{\text{SuccessRate}} = 0.03$, $t_{6374} = 4.80$, $p < 0.01$). That is, students who have the highest risk of not reading the feedback (i.e., being in the Assessment Results Only group) also tend to have higher test anxiety than students who read the feedback. They also tend to receive lower scores than students in the Assessment Results and Feedback group on the reading, writing and mathematics assessments at both institutions. The differences in scores are statistically significant, even with a Bonferroni adjustment (Table 7).

At EC, only the ATE for Retention using matching was statistically significant. At that institution, self-efficacy for online learning predicts Retention using propensity score weights (Beta$_{\text{_Retention}} = 0.16$, $t_{6381} = 2.06$, $p < 0.04$) when Assessment Results Only is compared to Control.

**Table 7**

*Differences in DAACS Academic Assessment Scores by Dosage*

| DAACS Academic Assessments | Results and Feedback | Results Only | *t*-statistic | *p*-value |
|---|---|---|---|---|
| | | Excelsior College | | |
| Mathematics | 0.60 (SD = 0.20) | 0.55 (SD = 0.20) | $t_{1960} = 6.53$ | < 0.01 |
| Reading | 0.87 (SD = 0.14) | 0.82 (SD = 0.20) | $t_{1639} = 6.08$ | < 0.01 |
| Writing | 0.82 (SD = 0.15) | 0.73 (SD = 0.17) | $t_{1481} = 11.39$ | < 0.01 |
| | | Western Governors University | | |
| Mathematics | 0.65 (SD = 0.17) | 0.58 (SD = 0.19) | $t_{5134} = 13.68$ | < 0.01 |
| Reading | 0.90 (SD = 0.10) | 0.86 (SD = 0.16) | $t_{3973} = 12.46$ | < 0.01 |
| Writing | 0.83 (SD = 0.14) | 0.75 (SD = 0.17) | $t_{4727} = 18.57$ | < 0.01 |

## Research Question 3: To What Extent Does the Inclusion of DAACS Results in Predictive Analytic Models Increase the Accuracy of Predictions of Academic Success?

To examine the predictive power of the DAACS data, the datasets from each institution were randomly split. Models were trained using 70% of the data and then validated with the remaining 30%. Prediction was assessed for our dichotomous outcome variable, term-to-term retention.

### Specifying Base and DAACS Models

Two models were estimated for each of the two dichotomous outcome measures for each institution, and model accuracy was determined using cross-validation: the initial, or base model with student information only, and the base model plus DAACS. In a review of literature on predicting academic success in higher education, Alyahyan and Düştegör (2020) found that student demographics and prior academic achievement were among the top two factors used to predict student performance. Accordingly, our base model included 10 demographic and academic background variables that were provided by the two institutions (Table 1): age, military, citizenship, employment status, first-generation student, gender, ethnicity, household income, current program, and transfer credits. The base model plus DAACS included results from the three academic DAACS measures (math, reading, writing) and the six scales of the SRL questionnaire (strategies, metacognition, self-efficacy, mindset, mastery orientation, anxiety). Dosage, in terms of number of DAACS page views and clicks to OERs, was also included as a predictor.

***Predictive Power of DAACS***

Confusion matrices and prediction accuracies are in Table 8. The improvement is the difference between the model accuracy and guessing. For example, the success rate at EC in the validation dataset is 56.3%. The simplest model would predict that all students succeed. Under this model, we would be accurate 56.3% of the time. The accuracy of the base model (the model including demographic variables) was 64.2%. For retention at EC, adding DAACS data to the predictive model increased accuracy by 7.8 percentage points, which is an increase of 4.0 percentage points over the base model. The level of improved prediction at WGU across retention was 0.1 percentage points. Thus, at both institutions, including DAACS data in the predictive models resulted in an increase in accuracy of between 0.1 and 4.0 percentage points over the base (demographic only) models.

**Table 8**

*Confusion Matrices and Prediction Accuracies for Base (demographics only) and Base Plus DAACS Models using Random Forests*

| | | | Base | | | Base plus DAACS | | |
|---|---|---|---|---|---|---|---|---|
| | | | Predicted | | | Predicted | | |
| Outcome | Observed | | Fail | Pass | | Fail | Pass | |
| Excelsior College | | | | | | | | |
| Retained | Fail | 61.9 | 49.8 | 12.1 | | 47.4 | 14.5 | |
| | Pass | 38.1 | 26.2 | 11.9 | | 19.8 | 18.3 | |
| | Accuracy | | | | 61.6 | | | 65.7 |
| | Improvement | | | | **-0.2** | | | **3.8 (4.0)** |
| Western Governors University | | | | | | | | |
| Retained | Fail | 26.7 | 1.2 | 25.5 | | 0.3 | 26.4 | |
| | Pass | 73.3 | 2.5 | 70.9 | | 0.3 | 73.0 | |
| | Accuracy | | | | 72.1 | | | 73.3 |
| | Improvement | | | | **-0.1** | | | **0.0 (0.1)** |

Note: Improvement is measured as the difference between model accuracy (sum of the main diagonal of the confusion matrix) and the larger observed success or failure proportions. The values in parentheses are the difference between the DAACS and base models. Values in bold are the overall difference between the base and DAACS predictive models.

## Discussion

This RCT on the effects of a field-based intervention had two purposes. First, we aimed to experimentally examine the effects of DAACS in enhancing the academic success of college students as measured by two critical outcomes at the college level: success rate and term-to-term retention. The second purpose, which has implications regarding institutional policy and curriculum, was to examine the additive effects of DAACS assessment data for predicting student performance. Treatment group differences were not evident at either EC or WGU when using an intent-to-treat approach (i.e., all students assigned to condition regardless of receiving the treatment). However, post hoc analyses revealed that students who received results from the four DAACS assessments and accessed the DAACS feedback did better in terms of success rate/credits earned and term-to-term retention than students who only accessed their results without reading the feedback. Post hoc analyses also suggest potentially negative effects of simply taking the assessments and receiving the results without also reading the feedback. This finding underscores how receiving assessment results may lead to negative effects in some contexts, while feedback could mitigate such effects.

This study also demonstrated the potential utility of DAACS assessment data from an administrator's or systemic perspective. That is, adding DAACS data (i.e., reading, mathematics, writing, and SRL information) to more traditional prediction models increased the accuracy of predicting students' early success in college across the two institutions.

**Effects of DAACS**

Although the central objective of this study was to examine the effects of the full implementation of DAACS in two college contexts, practical issues prevented the full complement of DAACS components from being universally implemented and adopted by students and their advisors. While disappointing and contradictory to our initial hypotheses, the observed null effects of DAACS on the two outcomes were not surprising, given that many students who were assigned to treatment condition at one of the participating institutions did not actually receive the intervention.

However, the post hoc PSA revealed that students who received DAACS assessment results and accessed the feedback outperformed those who chose not to access the feedback. These findings align with the feedback literature, which emphasizes the strong link between timely, informative feedback and improved academic outcomes (Bangert-Drowns et al., 1991; Grimaldi et al., 2019; Hattie & Timperley, 2007; Koenka et al., 2021).

The unexpected finding that receiving assessment results but not reading the feedback can lead to negative effects is also consistent with research on the effects of low grades or scores on student motivation and subsequent performance (Koenka et al., 2021). This is especially true for low-achieving students (Butler & Nisan, 1986) and those with low self-efficacy (Hattie & Timperley, 2007; Kim & Lee, 2019) or high test anxiety (von der Embse et al., 2018; Weissgerber & Reinhard, 2018). Although the DAACS system does not report grades, it does provide students with quasi-scores in the form of one, two, or three dots; student essays produced for the writing assessment include frequent references to those dots as "scores." One dot, which could be interpreted by students as a low score or failure, can provoke negative emotions. Emotional reactions are known to mediate the relationship between feedback and counterproductive behaviors such as disengagement (Belschak et al., 2009). Further, negative emotions can often lead to maladaptive outcomes, and are predictors of lower engagement and grades (Pekrun et al., 2023). That is, academic engagement is often a response to performance feedback, not just an antecedent of school performance (Poorthuis et al., 2015).

If students receive both a grade or score and narrative feedback, many of them will pay attention to the score and ignore the feedback (Brookhart, 2008; Bulut et al., 2019). We have some evidence that this occurred with the SRL questionnaire: Students were instructed to review the feedback on their SRL results before composing the essay for the writing assessment, but many did not click on the feedback. It appears that they wrote their essays based on the number of dots they received. In a stand-alone system like DAACS, students must be proactive recipients of feedback information. Because the positive treatment effect was statistically significant for the students who accessed the DAACS feedback, we speculate that viewing the feedback might mitigate the negative effects of having test anxiety or low self-efficacy and performing poorly on the assessments. This speculation is supported by research that shows that feedback received during an academic testing event can decrease feelings of worry and anxiety (Hong et al., 1999;

Morris & Fulmer, 1976). In addition, instructors and advisors can convey the message to low-performing students that their difficulties are temporary, thereby increasing feedback uptake (Winstone et al., 2021) and mitigating the negative effects of scores (Dweck, 2012; Poorthuis et al., 2015).

**Predictive Modeling**

The results of the predictive modeling support our claim that DAACS data can be of value at the institutional level. At a time when institutions are required to do more with less, being able to more accurately identify students most in need of support, along with detailed information about their individual needs, is encouraging. Moreover, the improvement in the accuracy of predictive models that include DAACS data indicates that it measures knowledge and skills that are associated with success in college. These results also provide evidence of the predictive validity of the assessments and dosage data.

Although predicting student success appears to be widely used by institutions, there is very little published research to which our results can compare. One study by Dalton et al. (2018) used high school transcript data to predict student success in first semester English with 67.1% accuracy, a 2.5% increase over the overall base success rate. Our results suggest that the use of cost-efficient, low resource, low-stakes assessment data, typically completed by students in under two hours, can provide as much or more predictive power than high school transcript data collected over the course of four years.

**Limitations and Future Research**

Although RCTs are considered the gold standard for establishing causality, there are limitations to analyses based only upon intent-to-treat (Imai et al., 2008). First, RCTs assume that the treatment effects are homogenous for all students. This study involves very diverse student groups with varying backgrounds in terms of educational experience, social economic status, and underrepresented populations. Second, RCTs assume that the intervention is uniformly received. Implementing an RCT at the institutional level involving all incoming students presented numerous challenges both anticipated and unexpected (see Cleary et al., 2023, for an overview of challenges and solutions during implementation). In fact, trace data from students' page views of feedback and links to OERs revealed that not all students received the same dosage of DAACS.

Another limitation of the study was the lack of full implementation of all DAACS components, particularly the use of college advisors to support students understanding and processing of DAACS feedback. According to the WWC guidelines, tutors, advisors, and instructional coaches can offer valuable help as students process and act upon DAACS feedback. Future research needs to bolster the role of the advising component of DAACS and to directly examine its effects on student outcomes. In addition, given research that indicates the superiority of feedback over grades (even when feedback accompanies grades; Lipnevich & Smith, 2008), the effects of reporting DAACS results without the quasi-scores (one, two or three dots) must be explored.

Another limitation was that our sample was largely non-traditional, adult learners with transfer credits enrolled in online colleges. Although targeting this underserved population and context is important, the degree to which the findings can be generalized to traditional, first-time college students in their late teens or other populations is limited. There are important differences between types of students. For example, most online students are full-time employees (Clinefelter & Aslanian, 2015), and non-traditional, adult students are more likely to be financially independent and have dependents such as a spouse or children (Gilardi &

Gugielmetti, 2011). While adult learners have more life and work experience than traditional college students, they will, on average, also encounter more life demands and challenges associated with parental roles and their vocations.

It is important that future research examine the effects of DAACS across diverse contexts and sample within the higher educational landscape. As pedagogical initiatives evolve over time, it is essential to examine whether DAACS has greater utility and effectiveness across online, hybrid or more traditional classroom contexts. Given the research on the disproportionately negative effects of placement exams on Black and Hispanic students (Park-Gaghan et al., 2020), it is important that future efficacy studies examine whether the effects of DAACS differ by race, as well as gender, socioeconomic status, and other demographic variables. Further, given the multi-component nature of DAACS (e.g., assessments, feedback, OERs, advising, instruction), the use of component analysis studies can be useful in identifying which individual or combination of DAACS components are most essential to producing changes in student success. Finally, the use of DAACS to support marginalized and diverse students, particularly first generation or immigrant students, is a high priority given the increasing number of students within such populations.

Further, as administrators continue to grapple with ways to incorporate cost-effective initiatives for improving student success and overall graduation rates, there is a need to go beyond the prediction algorithms discussed in this paper. Person-centered approaches to understanding student success in college, such as cluster analysis or latent profile analysis, are ideally suited to increase the focus on individual students rather than merely the importance of one variable relative to another in predicting outcomes (de Clercq et al., 2017; Woo et al., 2018).

Person-centered approaches have the potential to assist college instructors, advisors, and others who use DAACS in more accurately identifying students who are academically at-risk.

**Conclusion**

Overall, this study contributes to the literature regarding promising academic interventions among college populations and provides evidence that low-cost assessments of and feedback on academic skills and SRL upon enrollment in college may benefit both the individual student and the host institution. We believe DAACS has value in spite of the small positive effect sizes because light touch, low intensity interventions like DAACS typically produce smaller impacts than more intensive, expensive interventions (Jacob et al., 2019). Although DAACS was free to students and very inexpensive for the institutions, we provided evidence that its full implementation, in which students read the feedback and use the OERs to ensure they are prepared for college, is related to more positive student outcomes during the initial transition to college life. Jacob et al. argue that when an intervention demands a very low outlay of time and/or money, any impact makes the investment worthwhile.

To illustrate, consider that for the WGU sample of 13,200 students, the statistically significant but small 1.2% difference between the control condition and the assessments plus feedback treatment condition, applied across the entire sample, would mean that an additional 158 students would enroll in another term after taking the DAACS. Given WGU's current tuition rate of $6,670 per term, this translates to $1,053,860 of additional tuition paid to the University. Importantly, we would also predict retention of an additional 462 students at WGU if all 13,200 students had taken the DAACS. It appears that DAACS has the potential to enhance the academic and regulatory skills of a large number of college students without expensive and time-consuming interventions, including, perhaps, the aforementioned remediation courses that cost students approximately $7 billion per year. It is our expectation that when students, their advisors, and their institutions take advantage of the rich resources available to them through the DAACS system, more students can avoid remediation, take and pass credit-bearing courses, and complete their degrees.

**References**

ACT. (2019). *The Condition of College and Career Readiness 2019.* ACT. https://www.act.org/content/dam/act/secured/documents/cccr-2019/National-CCCR-2019.pdf.

Akhmedjanova, D., Lui, A. M., Andrade, H. L., & Bryer, J. (2019, April 4-8). *Validity and reliability of the DAACS writing assessment* [Paper presentation]. Annual meeting of the National Council on Measurement in Education (NCME), Toronto, Canada.

Alyahyan, E., Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education, 17*(3). https://doi.org/10.1186/s41239-020-0177-7

Andrade, H., Bryer, J., & Yagelski, R. (2018). *Developing and validating the DAACS writing assessment*. Paper presentation at the 16th international conference of the EARLI special interest group on writing, Antwerp, Belgium.

Andrade, H., & Cizek, G. (Eds.) (2010). *Handbook of formative assessment.* Routledge.

Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *The Journal of Higher Education, 77*(5), 886–924.

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching & Learning*, *4*(2). https://doi.org/10.20429/ijsotl.2010.040217

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. (2016). *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. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/wwc_dev_ed_112916.pdf

Bailey, T., & Cho, S-W. (2010). *Developmental education in community colleges* [CCRC Brief]. Community College Research Center, Columbia University. https://doi.org/10.7916/D8JS9NHX

Bandura, A. (2001). Social cognitive theory: An agentic perspective*. Annual Review of Psychology, 52*(1), 1-26.

Bangert-Drowns, R. L., Kulik, C.-L., Kulik J. L., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238. https://www.jstor.org/stable/1170535

Belschak, F. D., Den Hartog, D. N. (2009) Consequences of positive and negative feedback: The impact on emotions and extra-role behaviors. *Applied Psychology, 58*(2), 274-303. https://doi-org.libproxy.albany.edu/10.1111/j.1464-0597.2008.00336.x

Belfield, C. R., & Crosta, P. M. (2012). *Predicting success in college: The importance of placement tests and high school transcripts* [CCRC Working Paper No. 42]. Community College Research Center, Columbia University.

Boatman A. & Long B. T. (2018). Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation. *Educational Evaluation and Policy Analysis. 40*(1), 29-58. https://doi.org/10.3102/0162373717715708

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Harvard University Press.

Bryer, J., Andrade, H., Cleary, T., & Lui, A. (2018). *The diagnostic assessment and achievement of college skills*. Paper presented at the Distance Teaching and Learning Conference, Madison, WI.

Bulut, O., Cutumisu, M., Aquilina, A. M., & Singh, D. (2019). Effects of digital score reporting and feedback on students' learning in higher education. *Frontiers in Education, 4*. https://doi.org/10.3389/feduc.2019.00065

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 78*(3), 210-216. https://doi.org/10.1037/0022-0663.78.3.210

Cleary, T. (2006). The development and validation of the self-regulation strategy inventory—self-report. *Journal of School Psychology, 44*(4), 307–322. https://doi.org/10.1016/j.jsp.2006.05.002

Cleary, T. J., Bryer, J. M., & Andrade, H. A. (2023). Using the diagnostic assessment and achievement of college success (DAACS) to promote SRL skills among entry-level college students: Challenges and recommendations. *New Directions for Teaching and Learning, 2023*(4), 73-79. http://doi.org/10.1002/tl.20551

Clinefelter, D. L., & Aslanian, C. B. (2015). *Online college students 2015: Comprehensive data on demands and preferences*. The Learning House, Inc.

Colvard, N. B., Watson, C. E., & Park, H. (2018). The impact of open educational resources on various student success metrics. *International Journal of Teaching and Learning in Higher Education*, *30*(2), 262-276.

Cousert, D. (1999). The effects of a mentoring intervention program on retention of students in a community college (UMI No. 304550777) [Doctoral dissertation, Purdue University]. ProQuest Dissertations and Theses database.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences, 21*, 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Dalton, A., Beer, J., Kommanapalli, S., & Lanich, J.S. (2018). Machine learning to predict college course success. *SMU Data Science Review, 1*(2). https://scholar.smu.edu/cgi/viewcontent.cgi?article=1014&context=datasciencereview

de Brey, C., De La Rosa, J., Irwin, V., & Sonnenberg, W. C. (2022). Retention rate. In *National Center for Education Statistics*. NCES. https://nces.ed.gov/programs/coe/glossary#r

de Clercq, M., Galand, B. & Frenay, M. (2017). Transition from high school to university: a person-centered approach to academic achievement. *Eur J Psychol Educ, 32,* 39–59. https://doi.org/10.1007/s10212-016-0298-5

Driscoll, R. (2007). *Westside test anxiety scale validation*. ERIC Digest. https://files.eric.ed.gov/fulltext/ED495968.pdf

Dugan, R., & Andrade, H. L. (2011). Exploring the construct validity of academic self-regulation using a new self-report questionnaire–the survey of academic self-regulation. *International Journal of Educational and Psychological Assessment, 7*(1), 45-63.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House.

Dweck, C. S. (2012). *Mindset: How you can fulfill your potential.* Random House.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

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

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research, 15*(1), 3133-3181. https://www.jmlr.org/papers/volume15/delgado14a/delgado14a.pdf

Ganga, E., & Mazzariello, A. (2019). *Modernizing college course placement by using multiple measures*. Education Commission of the States. Center for the Analysis of Postsecondary Readiness (CAPR). https://www.ecs.org/wp-content/uploads/ Modernizing-College-Course-Placement-byUsing-Multiple-Measures.pdf

Gilardi, S., & Guglielmetti, C. (2011). University life of non-traditional students: Engagement styles and impact on attrition. *Journal of Higher Education*, *82*(1), 33-53.

Grimaldi, P. J., Basu Mallick, D., Waters, A. E., & Baraniuk, R. G. (2019). Do open educational resources improve student learning? Implications of the access hypothesis. *PLoS ONE 14*(3):e0212508. https://doi.org/10.1371/journal.pone.0212508

Grubb, W. N. (2001). *"Getting into the world": Guidance and counseling in community colleges*. Community College Research Center, Teachers College, Columbia University. https://ccrc.tc.columbia.edu/media/k2/attachments/getting-into-world-guidance-counseling.pdf

Han, N. (2003). MCAS 2001 grade 10 ELA and Mathematics model fit analyses (MCAS Validity Report No. 8. [CEA-540]). Center for Educational Assessment, University of Massachusetts. http://www.umass.edu/remp/docs/MCAS-RR-8.pdf

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J., & Timperley, H. (2007). *The power of feedback. Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Helmreich, J. E., & Pruzek, R. M. (2009). PSAgraphics: An R package to support propensity score analysis. *Journal of Statistical Software, 29*(6), 1-23. https://www.jstatsoft.org/v29/i06/.

Hilton, J. (2016). Open educational resources and college textbook choices: A review of research on efficacy and perceptions. *Educational Technology Research and Development, 64*(4), 573-590. https://doi.org/10.1007/s11423-016-9434-9

Hilton, J. (2020). Open educational resources, student efficacy, and user perceptions: a synthesis of research published between 2015 and 2018. *Education Technology Research Development, 68*, 853-876. https://doi.org/10.1007/s11423-019-09700-4

Hilton, J. L., Robinson, T. J., Wiley, D., & Ackerman, J. D. (2014). Cost-savings achieved in two semesters through the adoption of open educational resources. *International Review of Research in Open and Distributed Learning, 15*(2), 67-84. https://doi.org/10.19173/irrodl.v15i2.1700

Imai, K., King G., & Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 171*(2), 481-502. https://doi.org/10.1111/j.1467-985X.2007.00527.x

Jacob, R., Doolittle, F., Kemple, J., & Somers, M. (2019). A framework for learning from null results. *Educational Researcher, 48*(9), 580–589. https://doi.org/10.3102/0013189X19891955

Jaggars, S. S., & Stacey, G. W. (2014). *What we know about developmental education outcomes*. Community College Research Center, Teachers College, Columbia University. https://ccrc.tc.columbia.edu/media/k2/attachments/what-we-know-about-developmental-education-outcomes.pdf

Jirka, S. J., & Hambleton, R. K. (2005). Cognitive complexity levels for the MCAS assessments (MCAS Validity Report No. 10 [CEA-566]). Center for Educational Assessment, University of Massachusetts. https://www.umass.edu/remp/docs/MCAS-RR-10.pdf

Keele, L. (2010). An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. *White Paper. Columbus, OH*, *1*, 15.

Kim, E. J., & Lee, K. R. (2019). Effects of an examiner's positive and negative feedback on self-assessment of skill performance, emotional response, and self-efficacy in Korea: a quasi-experimental study. *BMC Medical Education, 19*(1), 142. https://doi.org/10.1186/s12909-019-1595-x

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement: a case for written feedback. *Educational Psychology, 41*(7), 922-947. https://doi.org/10.1080/01443410.2019.1659939

Leite, W. L. (2017). *Practical propensity score methods using R*. Thousand Oaks, CA: Sage Publishing.

Lipnevich, A. A., & Smith, J. K. (2008). Response to assessment feedback: The effects of grades, praise, and source of information. *ETS Research Report Series, 2008*(1), i-57. https://doi.org/10.1002/j.2333-8504.2008.tb02116.x

Lipnevich, A. & Smith, J. (Eds.). (2018). *Cambridge handbook of instructional feedback*. Cambridge University Press.

Lonjon, G., Boutron, I., Trinquart, L., Ahmad, N., Aim, F., Nizard, R., Ravaud, P. (2014). Comparison of treatment effect estimates from prospective nonrandomized studies with propensity score analysis and randomized controlled trials of surgical procedures. *Annals of Surgery*, *259*(1), 18-25. https://doi.org/10.1097/SLA.0000000000000256.

Lui, A., Franklin, D., Akhmedjanova, D., Gorgun, G., Bryer, J., Andrade, H., & Cleary, T. (2018). Validity evidence of the internal structure of the DAACS self-regulated learning survey. *Future Review: International Journal of Transition, College, and Career Success, 1*(1), 1-18. http://www.futureinstitute.us/wp-content/uploads/2019/03/Future-Review-Online-Article-1.1.pdf

Massachusetts Department of Elementary and Secondary Education (2017). 2016 MCAS and MCAS-Alt Technical Report. Measured Progress. http://mcasservicecenter.com/documents/MA/Technical%20Report/2016/2016%20MCAS%20%20MCAS%20Alt%20Technical%20Report_Final.pdf

Mayfield, E., & Rosé, C. P. (2013). LightSide: Open source machine learning for text. In M., Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluation (pp. 146-157)*. Routledge.

Meijer, E., Cleiren, M. P., Dusseldorp, E., Buurman, V. J., Hogervorst, R. M., & Heiser, W. J. (2019). Cross-validated prediction of academic performance of first-year university students: Identifying risk factors in a nonselective environment. *Educational Measurement: Issues and Practice*, *38*(1), 36-47.

Mokher, C.G., Barnett, E., Leeds, D.M., & Harris, J.C. (2019). Re-envisioning college readiness reforms: Florida's statewide initiative and promising practices in other states. *Change: The Magazine of Higher Learning, 52*(2), 14-23. https://doi.org/10.1080/00091383.2019.1569968

National Assessment of Educational Progress (2015). *The nation's report card: Mathematics and reading.* National Assessment of Educational Progress. https://www.nationsreportcard.gov/reading_math_2015/#?grade=4

National Center for Public Policy and Higher Education, & Southern Regional Education Board, (2010, June). *Beyond the rhetoric: Improving college readiness through coherent state policy (A Special Report).* (ERIC Document Reproduction Service No. ED 510 711). National Center for Public Policy and Higher Education. http://www.highereducation.org/reports/college_readiness/CollegeReadiness.pdf

Neal, J. W., & Neal, Z. P. (2013). Nested or networked? Future directions for ecological systems theory. *Social Development, 22*(4), 722-737. https://doi.org/10.1111/sode.12018

New York State Education Department (2014a). *New York State Regents examination in algebra 1 (Common Core): 2014 field test analysis, equating procedure, and scaling of operational test forms technical report.* Pearson. http://www.p12.nysed.gov/assessment/reports/commoncore/tr-a1-14.pdf

New York State Education Department (2014b). *New York State Regents examination in ELA (Common Core): 2014 field test analysis, equating procedure, and scaling of operational test forms technical report.* Pearson. http://www.p12.nysed.gov/assessment/reports/2014/hselacc-tr14.pdf

Park-Gaghan, T. J., Mokher, C. G., Hu, X., Spencer, H., & Hu, S. (2020). What happened following comprehensive developmental education reform in the sunshine state? The impact of Florida's developmental education reform on introductory college-level course completion. *Educational Researcher*, *49*(9), 656-666. https://doi.org/10.3102/0013189X20933876

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2023). School grades and students' emotions: Longitudinal models of within-person reciprocal effects. *Learning and Instruction, 83*. https://doi.org/10.1016/j.learninstruc.2022.101626

Poorthuis, A. M. G., Juvonen, J., Thomaes, S., Denissen, J. J. A., Orobio de Castro, B., & van Aken, M. A. G. (2015). Do grades shape students' school engagement? The psychological consequences of report card grades at the beginning of secondary school. *Journal of Educational Psychology, 107*(3), 842-854. https://doi.org/10.1037/edu0000002

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387.

Rodriguez, O., Bowden, B., Belfield, C., & Scott-Clayton, J. (2015, May). Calculating the costs of remedial placement testing. *Community College Research Center Analytics,* https://files.eric.ed.gov/fulltext/ED562055.pdf

Rosenbaum, P. R. 2002. Observational Studies. 2nd ed. New York, NY: Springer.

Rosenbaum, P. R. 2005. Observational study. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science: Vol. 3.* John Wiley and Sons.

Rosenbaum, P. R. (2012). Testing one hypothesis twice in observational studies. *Biometrika, 99*(4), 763-774. https://doi.org/10.1093/biomet/ass032

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475. http://wiki.biologyscholars.org/@api/deki/files/99/=Schraw1994.pdf

Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). Columbia University, Teachers College, Community College Research Center. https://ccrc.tc.columbia.edu/media/k2/attachments/high-stakes-predict-success.pdf

Scrivener, S., & Weiss, M. J. (2013, December). *More graduates: Two-year results from an evaluation of Accelerated Study in Associate Programs (ASAP) for developmental education students* [Policy Brief]. MDRC. https://www.mdrc.org/sites/default/files/More_Graduates.pdf

Sekhon, J. S. (2011). Multivariate and propensity score matching software with automated balance optimization: The matching package for R. *Journal of Statistical Software, 42*(7), 1-52. https://www.doi.org/10.18637/jss.v042.i07

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, *103*(484), 1334–1343. https://doi.org/10.1198/016214508000000733

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189. https://doi.org/10.3102/0034654307313795

Sinharay, S., Stern, H.S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6*(4), 317-329. https://doi.org/10.1037/1082-989X.6.4.317

State University of New York [SUNY] (2012). The SUNY pathway to success. The report of the SUNY task force on remediation. Office of the Senior Vice Chancellor Community Colleges and the Education Pipeline, SUNY. https://blog.suny.edu/wp-content/uploads/2013/01/Remediation-report-Jan-2.2013-complete.pdf

Stuart, B. L., Grebel, L. E., Butler, C. C., Hood, K., Verheij, T., & Little, P. (2017). Comparison between treatment effects in a randomised controlled trial and an observational study using propensity scores in primary care. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, *67*(662), e643–e649. https://doi.org/10.3399/bjgp17X692153

Van Buuren, S. & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1-67. https://www.jstatsoft.org/article/view/v045i03

Visher, M. G., Butcher, K. F., & Cerna, O. S. (2010). Guiding developmental math students to campus services: An impact evaluation of the Beacon Program at South Texas College. MDRC. https://www.mdrc.org/sites/default/files/full_382.pdf

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483-493.

Weinstein, C. E., Palmer, D. R., & Schulte, A. (1998). *The learning and study strategies inventory, high school edition.* H & H Publishing.

Weissgerber, S. C., & Reinhard, M. A. (2018). Pilot study on the relationship of test anxiety to utilizing self-testing in self-regulated learning. *International Journal of Psychological Studies, 10*(4), 95-109. doi:10.5539/ijps.v10n4p95

What Works Clearinghouse (2017). WWC procedures and standards handbook (version 4.0). Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Wiliam, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning (pp. 53–82)*. Lawrence Erlbaum.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice (pp. 277-304)*. Lawrence Erlbaum.

Winstone, N., Bourne, J., Medland, E., Niculescu, I., & Rees, R. (2021). "Check the grade, log out": students' engagement with feedback in learning management systems. *Assessment & Evaluation in Higher Education, 46*(4), 631-643.

Wolters, C. A., & Hoops, L. D. (2015). Self-regulated learning interventions for motivationally disengaged college students. In T. J. Cleary (Ed.), *School psychology series. Self-regulated learning interventions with at-risk youth: Enhancing adaptability, performance, and wellbeing* (pp. 67–88). American Psychological Association.

Woo, S. E., Jebb, A. T., Tay, L., & Parrigon, S. (2018). Putting the “person” in the center: review and synthesis of person-centered approaches and methods in organizational science. *Organizational Research Methods, 2*(4), 814-845. https://doi.org/10.1177/1094428117752467

Xu, Y., & Goodacre, R. (2018). On splitting training and validation set: A comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of Analysis and testing, 2*, 249-262. https://doi.org/10.1007/s41664-018-0068-2

Yagelski, R. (2015). *The essentials of writing: Ten core concepts*. Cengage Learning.

Yu, E-C., Rosca, O., & Andrade, H. (2022, August 4-6). *Predictive validity evidence for machine-scored diagnostic writing assessment* [Poster presentation]. American Psychological Association Annual Convention, Minneapolis, MN.

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and overview. In B. J Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance: Educational psychology handbook*. Routledge.

Zimmerman, B. J., Moylan, A., Hudesman, J., White, N., & Flugman, B. (2011). Enhancing self-reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling, 53*(1), 141-160.

Zimmerman, B.J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Self-regulation: Theory, research, and applications* (pp. 13-39). Academic Press.

Zhadko, O., & Ko, S. (2019). *Best practices in designing courses with open educational resources.* Taylor & Francis.

**Appendix A**

*Four DAACS Assessment Domains, Sample Feedback and Resources*

| *Domain* | *Sample Feedback* | *Sample Resources* |
|---|---|---|
| Self-regulated learning | Metacognition (Planning): Your responses to the SRL assessment indicate that you sometimes plan how to approach your schoolwork. Thus, before completing an assignment you might think about the most important parts of the assignment, or the best approach needed to complete it. Because planning can help students focus their attention and become more proficient in what they do, you may want to consider using the following planning strategies:<br>• Ask yourself questions before you begin a learning activity: "What am I expected to do? What approach can help me do well?"<br>• Brainstorm multiple ways to approach an activity and then choose the best option.<br>• Ask your teachers about the best ways to go about doing their assignments. | Self-Regulated Learning Lab: https://srl.daacs.net/ |
| Writing | Organization (Transitions): "A transition is a word or phrase that helps your reader follow your discussion easily from one paragraph or section of your essay to the next by making the logic and organization of your ideas clear. Transitions often include specific terms that indicate a sequence, such as "Another strength in my SRL is…" or "The second way I will improve is…". Explicit transitions might also be made through the use of specialized terms that indicate logical or semantic connections between statements or paragraphs, such as "As a result," "Therefore," or "By contrast." Sometimes the connection between one paragraph and another is indicated by the subject matter or repeated words rather than by explicit transitional terms. Whether you use explicit transitional words or not, transitions are necessary when there is a shift in the focus or topic from one paragraph to the next, and when moving from one main section of a document to another.<br>Your score suggests that the paragraphs in your essay are usually linked with transitions, as needed, but could be clearer or more effective. The transitions might be implied or strained, but a reader can still generally follow your discussion. If you would like to learn how to use transitions more effectively, we recommend following this link: https://owl.excelsior.edu/writing-process/paragraphing/paragraphing-and-transitioning/ | Excelsior College's Online Writing Lab: https://owl.excelsior.edu |
| Mathematics | Statistics: Showing data visually using scatterplots, line graphs, frequency distributions, and tables, is an important set of tools to look at and understand the data. You can understand other features of data using numeric summaries, like measures of central tendency (mean, median, and mode), correlation, and lines of best fit. All these tools are only helpful if you know how to read and interpret them. And since all these statistical tools are only as good | Math is Fun: https://www.mathsisfun.com/<br><br>Math in Action: http://www.mathinaction.org/ |

| *Domain* | *Sample Feedback* | *Sample Resources* |
|---|---|---|
| | as the data that is used to produce them, you need to know the questions to ask about how data was collected.<br>Your results suggest that your skills are just developing for reasoning with data. To further develop your skills at summarizing data with statistics, graphs, and tables, the resources below are a good starting point.<br>• Mean, median, and mode: http://www.mathsisfun.com/data/central-measures.html<br>• Graphs to display data:<br>    o Histograms: http://www.mathsisfun.com/data/histograms.html<br>    o Dot plots: http://www.mathsisfun.com/data/dot-plots.html<br>    o Scatterplots and lines of best fit: http://www.mathsisfun.com/data/scatter-xy-plots.html<br>• Conducting surveys to get reliable data: http://www.mathsisfun.com/data/survey-conducting.html | |
| Reading | Inference: Many texts convey information using indirect or suggestive language. For instance, it's common for authors to imply their thesis (main argument), or to require the reader to "read between the lines" to understand what is being communicated. This is particularly common in literary texts, advertising media, and persuasive texts. Therefore, it is important to be able to infer (make an educated guess based on evidence and logic) what the text is saying. Failure to infer meaning, or to make faulty inferences, may lead to misunderstanding.<br>Your results suggest an area of improvement for you is reading closely to determine implied meaning. A skill you may want to improve is the ability to draw logical inferences from what texts say explicitly to determine the implied meaning. An inference is an educated guess about the implied meaning of a passage or text based on a combination of evidence from the text and reasoning. You might consider using the following strategies to improve your inferencing skills:<br>• Use signposts to contextualize what you read and help guide your inferences<br>• Ask questions before, during, and after reading to improve the accuracy of your inferences<br>• Be aware of logical fallacies and personal biases that can lead to faulty reasoning and false inferences<br>• Analyze what a text is trying to say or imply by breaking it down into its major parts<br>• Annotate texts as you read to improve your overall reading comprehension | Reading Comprehension Lab: https://owl.excelsior.edu/orc/ |

Appendix B

Criterion Validity Evidence for the Reading and Mathematics Assessments: Full Matrix

| | Intercorrelations among Mathematics and Reading Subscales and Total Scores | | | | | | | | | | | | | Correlations with On-Time Progress | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | EC | WGU |
| Structure (1) | | 0.4 | 0.26 | 0.33 | 0.48 | 0.71 | 0.24 | 0.17 | 0.19 | 0.22 | 0.23 | 0.25 | 0.33 | 0.00 | 0.16 |
| Inference (2) | 0.29 | | 0.31 | 0.35 | 0.51 | 0.69 | 0.24 | 0.12 | 0.16 | 0.24 | 0.2 | 0.24 | 0.3 | -0.01 | 0.13 |
| Language (3) | 0.2 | 0.23 | | 0.26 | 0.41 | 0.57 | 0.22 | 0.09 | 0.15 | 0.18 | 0.18 | 0.21 | 0.26 | -0.04 | 0.10 |
| Purpose (4) | 0.23 | 0.27 | 0.2 | | 0.48 | 0.62 | 0.19 | 0.1 | 0.13 | 0.19 | 0.17 | 0.21 | 0.25 | -0.02 | 0.09 |
| Ideas (5) | 0.4 | 0.43 | 0.32 | 0.36 | | 0.86 | 0.31 | 0.16 | 0.19 | 0.3 | 0.27 | 0.29 | 0.38 | -0.03 | 0.17 |
| Read Total (6) | 0.68 | 0.62 | 0.52 | 0.53 | 0.82 | | 0.35 | 0.18 | 0.24 | 0.33 | 0.3 | 0.34 | 0.44 | -0.02 | 0.21 |
| Word Problems (7) | 0.2 | 0.16 | 0.14 | 0.1 | 0.25 | 0.28 | | 0.27 | 0.37 | 0.43 | 0.32 | 0.37 | 0.69 | 0.01 | 0.14 |
| Lines/Functions (8) | 0.12 | 0.08 | 0.09 | 0.07 | 0.12 | 0.15 | 0.24 | | 0.36 | 0.28 | 0.21 | 0.27 | 0.61 | 0.02 | 0.06 |
| Vars/Equations (9) | 0.15 | 0.11 | 0.11 | 0.08 | 0.16 | 0.2 | 0.28 | 0.28 | | 0.42 | 0.26 | 0.34 | 0.7 | -0.01 | 0.08 |
| Num/Calc (10) | 0.19 | 0.17 | 0.15 | 0.12 | 0.22 | 0.26 | 0.36 | 0.26 | 0.34 | | 0.28 | 0.38 | 0.7 | -0.01 | 0.11 |
| Statistics (11) | 0.18 | 0.14 | 0.12 | 0.1 | 0.19 | 0.23 | 0.26 | 0.15 | 0.21 | 0.24 | | 0.3 | 0.59 | 0.01 | 0.13 |
| Geometry (12) | 0.2 | 0.15 | 0.14 | 0.11 | 0.23 | 0.26 | 0.29 | 0.2 | 0.28 | 0.3 | 0.26 | | 0.67 | 0.01 | 0.14 |
| Math Total (13) | 0.27 | 0.22 | 0.2 | 0.15 | 0.31 | 0.37 | 0.65 | 0.57 | 0.65 | 0.66 | 0.57 | 0.63 | | 0.01 | 0.18 |

***Note:*** Correlations for EC are above the diagonal; for WGU below the diagonal. Dark grey = correlations between Math and Reading at each institution. Light grey = correlations between the domain (math or reading) and its corresponding subdomains Medium grey = correlations between the domain and subdomains of the other domain (e.g., math to reading subdomains). Correlations between (sub)domain scores and on-time progress are in the right-most columns.

**Appendix C**

Logistic Regression to Estimate Propensity Scores

| | Excelsior College | | | | | Western Governors University | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Effect** | **Estimate** | **SE** | ***z*** | ***p*-value** | | **Estimate** | **SE** | ***z*** | ***p*-value** | |
| (Intercept) | 0.20 | 0.28 | 0.73 | 0.46 | | -0.92 | 0.16 | -5.86 | 0.00 | *** |
| Age | 0.01 | 0.00 | 2.49 | 0.01 | * | 0.04 | 0.00 | 12.61 | 0.00 | *** |
| Military Student (Y) | -0.36 | 0.10 | -3.51 | 0.00 | *** | -0.36 | 0.09 | -4.06 | 0.00 | *** |
| English Language Native | -0.29 | 0.16 | -1.80 | 0.07 | . | -- | -- | -- | -- | |
| Citizenship (Non-Citizen) | -- | -- | -- | -- | | 0.26 | 0.29 | 0.92 | 0.36 | |
| Citizenship (Non-Resident Alien) | -- | -- | -- | -- | | 0.67 | 0.42 | 1.59 | 0.11 | |
| Employment (Part-time) | -0.14 | 0.18 | -0.77 | 0.44 | | 0.04 | 0.10 | 0.39 | 0.70 | |
| Employment (Full-time) | -0.13 | 0.11 | -1.24 | 0.21 | | -0.20 | 0.08 | -2.42 | 0.02 | * |
| First-Generation (Y) | 0.06 | 0.09 | 0.72 | 0.47 | | -0.06 | 0.05 | -1.18 | 0.24 | |
| Gender (Female) | -0.05 | 0.10 | -0.55 | 0.59 | | 0.31 | 0.07 | 4.68 | 0.00 | *** |
| Ethnicity (Am. Indian/Alaskan Native) | 0.68 | 0.48 | 1.44 | 0.15 | | -0.06 | 0.29 | -0.20 | 0.84 | |
| Ethnicity (Asian) | 0.28 | 0.21 | 1.32 | 0.19 | | 0.52 | 0.17 | 3.07 | 0.00 | ** |
| Ethnicity (Black or African Am.) | -0.28 | 0.11 | -2.59 | 0.01 | ** | -0.09 | 0.08 | -1.06 | 0.29 | |
| Ethnicity (Hispanic) | -0.04 | 0.12 | -0.32 | 0.75 | | 0.18 | 0.15 | 1.16 | 0.24 | |
| Ethnicity (Native Hawaiian/Other Pacific Islander) | -0.09 | 0.41 | -0.23 | 0.82 | | -0.18 | 0.30 | -0.59 | 0.56 | |
| Ethnicity (Multiple) | 0.28 | 0.18 | 1.53 | 0.13 | | 0.36 | 0.14 | 2.57 | 0.01 | * |
| Ethnicity (Unknown) | -0.07 | 0.39 | -0.19 | 0.85 | | -- | -- | -- | -- | |
| Income | -0.01 | 0.02 | -0.31 | 0.76 | | 0.04 | 0.02 | 2.20 | 0.03 | * |
| Transfer Credits | 0.00 | 0.00 | -2.21 | 0.03 | * | -0.01 | 0.00 | -5.64 | 0.00 | *** |
| Division (Business) | 0.02 | 0.12 | 0.15 | 0.88 | | -- | -- | -- | -- | |
| Division (Health Sciences) | 0.17 | 0.17 | 0.99 | 0.32 | | -- | -- | -- | -- | |
| Division (Nursing) | -0.15 | 0.14 | -1.03 | 0.30 | | -- | -- | -- | -- | |
| Division (Public Services) | 0.06 | 0.16 | 0.41 | 0.68 | | -- | -- | -- | -- | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Division (Technology) | 0.27 | 0.11 | 2.34 | 0.02 * | -- | -- | -- | -- |
| WGU Programs | | | | | | | | |
| Interdisciplinary Studies | -- | -- | -- | -- | -0.11 | 0.09 | -1.24 | 0.22 |
| Accounting | -- | -- | -- | -- | 0.06 | 0.10 | 0.57 | 0.57 |
| Healthcare Management | -- | -- | -- | -- | -0.18 | 0.10 | -1.78 | 0.07 . |
| Human Resource Management | -- | -- | -- | -- | -0.06 | 0.12 | -0.55 | 0.58 |
| IT Management | -- | -- | -- | -- | 0.17 | 0.14 | 1.25 | 0.21 |
| Science, Cloud & Systems Admin | -- | -- | -- | -- | 0.37 | 0.24 | 1.54 | 0.12 |
| Cybersecurity & Info Assurance | -- | -- | -- | -- | -0.14 | 0.30 | -0.48 | 0.63 |
| Data Management & Analysis | -- | -- | -- | -- | 0.18 | 0.20 | 0.91 | 0.37 |
| Health Information Management | -- | -- | -- | -- | 0.19 | 0.20 | 0.93 | 0.35 |
| Information Technology (IT) | -- | -- | -- | -- | 0.30 | 0.16 | 1.94 | 0.05 . |
| IT Network Admin | -- | -- | -- | -- | 0.36 | 0.18 | 2.01 | 0.04 * |
| IT Security | -- | -- | -- | -- | 0.25 | 0.14 | 1.73 | 0.08 . |
| IT Software Development | -- | -- | -- | -- | 0.35 | 0.14 | 2.46 | 0.01 * |
| Marketing | -- | -- | -- | -- | -0.12 | 0.14 | -0.84 | 0.40 |
| Network Operations & Security | -- | -- | -- | -- | 0.23 | 0.21 | 1.06 | 0.29 |
| Other Programs | -- | -- | -- | -- | 0.08 | 0.23 | 0.34 | 0.73 |
| Advisor Used (Y) | -- | -- | -- | -- | 0.05 | 0.06 | 0.86 | 0.39 |

*$p < .05$; **$p < .01$; ***$p < .001$

*Note:* Reference Groups: Employment = *Unemployed*; First-Generation = *N*; Ethnicity = *White*; Gender = *Male*; Military Student = *N*; Citizenship = *U.S. Citizen*; Advisor Used = N; Current Program = *BSMG*

**Appendix D**

Balance of Covariates Before (Unadjusted) and After (Adjusted) Propensity Score Matching at EC and WGU

Balance of Covariates at Excelsior College (EC)

| Covariate | Assessment + Feedback vs. Assessment Only | | Assessment + Feedback vs. Control | |
|---|---|---|---|---|
| | Unadjusted | Adjusted | Unadjusted | Adjusted |
| Age | 0.05 | 0.01 | 0.00 | 0.02 |
| DIVISION_CODE_BU | 0.02 | 0.00 | 0.01 | 0.01 |
| DIVISION_CODE_HS | 0.04 | 0.00 | 0.02 | 0.00 |
| DIVISION_CODE_LA | 0.06 | 0.01 | 0.02 | 0.01 |
| DIVISION_CODE_NU | 0.01 | 0.01 | 0.06 | 0.05 |
| DIVISION_CODE_PS | 0.01 | 0.00 | 0.05 | 0.05 |
| DIVISION_CODE_TE | 0.05 | 0.00 | 0.01 | 0.03 |
| EMPLOYMENT_STATUS_Full-Time | 0.04 | 0.00 | 0.02 | 0.00 |
| EMPLOYMENT_STATUS_Part-Time | 0.01 | 0.00 | 0.04 | 0.04 |
| EMPLOYMENT_STATUS_Unemployed | 0.04 | 0.01 | 0.01 | 0.02 |
| ENGLISH_LANGUAGE_NATIVE | 0.09 | 0.02 | 0.06 | 0.03 |
| ETHNICITY_American Indian or Alaska Native | 0.06 | 0.04 | 0.04 | 0.03 |
| ETHNICITY_Asian | 0.07 | 0.02 | 0.13 | 0.10 |
| ETHNICITY_Black or African American | 0.06 | 0.00 | 0.08 | 0.05 |
| ETHNICITY_Hispanic | 0.03 | 0.00 | 0.04 | 0.03 |
| ETHNICITY_Native Hawaiian or Other Pacific Islander | 0.01 | 0.00 | 0.02 | 0.02 |
| ETHNICITY_Two or more races | 0.04 | 0.00 | 0.03 | 0.00 |
| ETHNICITY_Unknown | 0.02 | 0.00 | 0.00 | 0.01 |
| ETHNICITY_White | 0.03 | 0.01 | 0.04 | 0.03 |
| First_Generation | 0.03 | 0.01 | 0.00 | 0.02 |
| GENDER_FEMALE | 0.03 | 0.01 | 0.01 | 0.00 |
| INCOME_RANGE_CODE | 0.01 | 0.00 | 0.05 | 0.04 |
| Initial_TRANSFER_CREDITS_EARNED | 0.09 | 0.00 | 0.06 | 0.02 |
| MilitaryStudent_TRUE | 0.12 | 0.00 | 0.07 | 0.01 |

Balance of Covariates at Western Governors University (WGU)

| Covariate | Assessment + Feedback vs. Assessment Only | | Assessment + Feedback vs. Control | |
|---|---|---|---|---|
| | Unadjusted | Adjusted | Unadjusted | Adjusted |
| Age | 0.35 | 0.01 | 0.10 | 0.00 |
| CITIZENSHIP_STATUS_Non-Citizen | 0.03 | 0.01 | 0.04 | 0.05 |
| CITIZENSHIP_STATUS_Non-Resident Alien | 0.05 | 0.02 | 0.01 | 0.00 |
| CITIZENSHIP_STATUS_U.S. Citizen | 0.05 | 0.02 | 0.03 | 0.04 |
| CURRENT_PROGRAM_CODE_BAISK8 | 0.00 | 0.00 | 0.02 | 0.01 |
| CURRENT_PROGRAM_CODE_BSAC | 0.02 | 0.00 | 0.00 | 0.01 |
| CURRENT_PROGRAM_CODE_BSBUHCM | 0.02 | 0.00 | 0.01 | 0.00 |
| CURRENT_PROGRAM_CODE_BSBUHR | 0.00 | 0.00 | 0.02 | 0.02 |
| CURRENT_PROGRAM_CODE_BSBUITM | 0.03 | 0.00 | 0.01 | 0.00 |
| CURRENT_PROGRAM_CODE_BSCLSA | 0.01 | 0.01 | 0.01 | 0.01 |
| CURRENT_PROGRAM_CODE_BSCSIA | 0.04 | 0.00 | 0.05 | 0.04 |
| CURRENT_PROGRAM_CODE_BSDMDA | 0.03 | 0.01 | 0.00 | 0.01 |
| CURRENT_PROGRAM_CODE_BSHIM | 0.06 | 0.01 | 0.03 | 0.01 |
| CURRENT_PROGRAM_CODE_BSIT | 0.03 | 0.00 | 0.03 | 0.02 |
| CURRENT_PROGRAM_CODE_BSITNW | 0.01 | 0.00 | 0.03 | 0.03 |
| CURRENT_PROGRAM_CODE_BSITSEC | 0.04 | 0.00 | 0.02 | 0.03 |
| CURRENT_PROGRAM_CODE_BSITSW | 0.03 | 0.00 | 0.01 | 0.01 |
| CURRENT_PROGRAM_CODE_BSMG | 0.01 | 0.00 | 0.01 | 0.01 |
| CURRENT_PROGRAM_CODE_BSMK | 0.05 | 0.00 | 0.02 | 0.01 |
| CURRENT_PROGRAM_CODE_BSNOS | 0.01 | 0.00 | 0.07 | 0.07 |
| CURRENT_PROGRAM_CODE_Other | 0.00 | 0.01 | 0.00 | 0.00 |
| EMPLOYMENT_STATUS_Full-Time | 0.05 | 0.01 | 0.00 | 0.02 |
| EMPLOYMENT_STATUS_Part-Time | 0.02 | 0.00 | 0.02 | 0.02 |
| EMPLOYMENT_STATUS_Unemployed | 0.05 | 0.01 | 0.02 | 0.00 |
| ETHNICITY2_Am. Indian or Alaskan Native | 0.01 | 0.00 | 0.02 | 0.02 |
| ETHNICITY2_Asian | 0.08 | 0.01 | 0.03 | 0.01 |
| ETHNICITY2_Black | 0.04 | 0.01 | 0.01 | 0.02 |
| ETHNICITY2_Hispanic | 0.02 | 0.01 | 0.05 | 0.05 |

| | | | | |
|---|---|---|---|---|
| ETHNICITY2_Multiple | 0.06 | 0.00 | 0.00 | 0.02 |
| ETHNICITY2_Native Hawaiian | 0.01 | 0.00 | 0.00 | 0.01 |
| ETHNICITY2_White | 0.03 | 0.00 | 0.04 | 0.03 |
| FIRST_GEN_STUDENT_TRUE | 0.01 | 0.00 | 0.01 | 0.00 |
| GENDER_Female | 0.12 | 0.01 | 0.03 | 0.00 |
| HOUSEHOLD_INCOME | 0.11 | 0.02 | 0.03 | 0.01 |
| MilitaryStudent_TRUE | 0.12 | 0.01 | 0.05 | 0.02 |
| TRANSFER_CREDITS | 0.13 | 0.02 | 0.05 | 0.02 |