



Online Credit Recovery Study

Technical Supplement for Research Briefs 1 – 4

MAY 2021

Iliana Brodziak de los Reyes | Jordan Rickles | Drew Atchison | Mark Lachowicz | Linda Lin | Peggy Clements | Emily Collins

MAKING RESEARCH RELEVANT

Online Credit Recovery Study

Technical Supplement for Research Briefs 1 – 4

MAY 2021

Iliana Brodziak de los Reyes | Jordan Rickles | Drew Atchison | Mark Lachowicz | Linda Lin | Peggy Clements | Emily Collins



AMERICAN INSTITUTES FOR RESEARCH®

1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
202.403.5000

www.air.org

Notice of Trademark: “American Institutes for Research” and “AIR” are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.

Copyright © 2021 American Institutes for Research®. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on www.air.org.

Contents

	Page
Introduction	1
Study Sample	1
Data and Measures	6
Analysis Approach	12
Analytic Models	13
Missing Data Approach	14
Impact Results	15
Main Impact Estimates	15
Results From Moderator Analysis	16
Results From Sensitivity Analysis	19
Cost Analysis	22
Cost Methodology	22
References	26
Appendix 1. Theory of Action	27

Tables

	Page
Table 1. Description of the Student Sample	3
Table 2. Comparison of the Total Student Sample to the Sample of Test Takers for Algebra 1, by Treatment Group	4
Table 3. Comparison of the Total Student Sample to the Sample of Test Takers for English 9, by Treatment Group	5
Table 4. Teacher Survey Questions for Each Instructional Feature.....	8
Table 5. Student Survey Questions for Each Instructional Experience	10
Table 6. Response Rates for the Student Survey and Test	14
Table 7. Estimated Effect on Instructional Features as Measured by the Teacher Survey	15
Table 8. Estimated Effect on Student Experiences as Measured by the Student Survey.....	15
Table 9. Estimated Effect on Student Test Scores	16
Table 10. Estimated Effect on Student Credit Recovery Rate	16
Table 11. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Experiences	16
Table 12. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Test Scores	17
Table 13. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Credit Recovery Rate.....	17
Table 14. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Experiences	17
Table 15. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Test Scores	18
Table 16. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Credit Recovery Rate	18
Table 17. Estimated Interaction Effect of Treatment and Eighth-Grade English Assessment on Student Experiences (English 9 Classes Only).....	18
Table 18. Estimated Interaction Effect of Treatment and Eighth-Grade English Assessment on Student Test Scores and Student Credit Recovery Pass Rate (English 9 Classes Only)	18

Table 19. Estimated Interaction Effect of Treatment and Eighth-Grade Math Assessment on Student Experiences (Algebra 1 Classes Only)..... 19

Table 20. Estimated Interaction Effect of Treatment and Eighth-Grade Math Assessment on Student Test Scores and Student Credit Recovery Pass Rate (Algebra 1 Classes Only) 19

Table 21. Estimated Effect on Student Experiences as Measured by the Student Survey, Without Imputation of Missing Data 20

Table 22. Estimated Effect on Student Test Scores, Without Imputation of Missing Data..... 20

Table 23. Sensitivity Analysis Estimated Effect on Student Experiences as Measured by the Student Survey for Schools That Complied With Randomization Procedure 21

Table 24. Sensitivity Analysis Estimated Effect on Student Test Scores for Schools That Complied With Randomization Procedure 21

Table 25. Sensitivity Analysis Estimated Effect on Student Credit Recovery Rate for Schools That Complied With Randomization Procedure..... 21

Table 26. Description of Resources Accounted for in the Cost Analysis and Their Categorization 24

Figure 1. Theory of Action..... 28

Acknowledgments: This study is funded by research grant R305A170152 from the U.S. Department of Education’s Institute of Education Sciences. The opinions expressed herein are those of the authors and do not represent the views of the Institute of Education Sciences, the U.S. Department of Education, or the Los Angeles Unified School District. The authors would like to acknowledge the many people who helped make this study possible. We thank the Los Angeles Unified School District leaders who worked with us on this study, particularly Carol Alexander; the staff at Edgenuity, who supported the participating schools, particularly Lindsay Marczak; and the school leaders, teachers, and students who participated in the study.

Introduction

To expand our understanding of the effectiveness of online credit recovery programs, the American Institutes for Research® (AIR®) conducted a multisite randomized study that tested an online learning model for credit recovery in Los Angeles high schools. The primary curriculum for the online learning model was provided through an online program (Edgenuity), and an in-class teacher provided additional individualized instructional support. The primary analyses focused on the implementation and initial outcomes of two cohorts of students who enrolled in an Algebra 1 or English 9 credit recovery course the summer following their first year of high school.¹ In addition to the summer sessions, the study included credit recovery classes held during the 2018–2019 school year, both online and in teacher-directed classroom settings. However, our analyses and research briefs focus on the summer sessions,² because there were meaningful qualitative differences in the implementation context between the school-year and summer classes, and the relatively small school-year sample limits the inferences we could make about the school year. In future work we will explore differences between the summer and school-year credit recovery classes.

This technical supplement provides a description of the study’s sampling procedures, data sources, and analytic methods for the findings presented in the study’s [first four research briefs](#).

Study Sample

In this section, we provide an overview of the study sample and student selection.

To be included in the study, students had to meet the following criteria:

- Entered ninth grade in the 2017–2018 or 2018–2019 school-year (expected graduation class of 2021 or 2022).
- Enrolled in a district high school in spring 2018 (for class of 2021) or spring 2019 (for class of 2022).
- Failed Algebra 1 and/or at least one semester of their English 9 course during their first year of high school.

¹ The study targeted two ninth-grade courses with high failure rates in the school district. Students in the school district typically take a yearlong first-year Algebra course (Algebra 1) in ninth grade and two semester-long English courses. English 9A is typically taken in the fall and English 9B is typically taken in the spring. We included English 9A and 9B classes in the study, and in this paper, we report on them together as English 9.

² There was only one session of credit recovery classes in each of the summer, fall, and spring periods.

- Enrolled in one of the study credit recovery classes (Algebra 1, English 9A, or English 9B) at the start of the term. For the first summer term (2018), only English 9 classes were in the study. For the second summer term (2019), the study included both Algebra 1 and English 9 classes.
- Were not classified with an English language development (ELD) level of 1, 2, or 3.³

The analyses in the first four research briefs are based on 613 students in 28 Algebra 1 classes across 13 high schools, and 1,124 students in 70 English 9 classes across 19 high schools.⁴ In each participating school, half of the classes were online classes and half were teacher-directed classes. We recruited traditional high schools in the Los Angeles Unified School District (LAUSD) that had a relatively large number of students who failed Algebra 1 and/or English 9 during the 2016–17 school year. We targeted schools with enough eligible students to support at least two credit recovery classes for a given course.

Students were randomly assigned to take their credit recovery course in an online class (treatment) or a teacher-directed class (control). Random assignment took place within blocks defined by subject, cohort, and school. In some schools, blocks were further defined by which semester of the course the students failed during their ninth-grade year.⁵

Table 1 compares the baseline characteristics of the students in the online and teacher-directed classes. As expected, for students requiring credit recovery, our sample performed well below average in eighth-grade English and math and had lower than a C average in their ninth-grade courses. Students in the online and teacher-directed classes had similar characteristics and prior academic struggles, on average, with standardized mean differences (SMD) less than 0.25 standard deviations (a common threshold for baseline equivalence). This indicates that the random assignment process successfully resulted in having similar students in each class type.

Not all students in the study sample completed the student survey and study-administered test (described in the next section); the response rate was 59% for English 9 and 72% for Algebra 1. Tables 2 and 3 compare the baseline characteristics of the total student sample with the test takers for the online and teacher-directed classes for Algebra 1 and English 9, respectively.

³ English learners are classified into one of five ELD levels, where a higher number indicates more advanced English language development. Per district policy, students with an ELD level below level 4 should not be enrolled in online courses, so we excluded them from the study.

⁴ In regard to teachers, in most schools one teacher taught both the online class and the teacher-directed class.

⁵ Rather than use the random assignment list we generated for student enrollment (which included blocking based on semesters failed), some schools elected to use a simplified version of random assignment based on whether the last digit of the student's district identification (ID) number was an even or odd number. After confirming that the last digit of the ID essentially functions as a random number, we allowed schools to use this option to facilitate school participation in the study.

Table 1. Description of the Student Sample

Student characteristics	Algebra 1			English 9		
	Online classes	Teacher-directed classes	SMD	Online classes	Teacher-directed classes	SMD
Number of students	305	308		564	560	
Female	47%	43%	0.08	35%	33%	0.04
Ethnicity: African American/Black	10%	12%	-0.09	8%	9%	-0.06
Ethnicity: Latinx/Hispanic	81%	80%	0.02	85%	83%	0.09
Ethnicity: other	9%	8%	0.08	7%	9%	-0.11
National school lunch–eligible	80%	80%	0.02	89%	90%	-0.04
Gifted/talented program	6%	8%	-0.18	12%	12%	0.02
Student with a disability	10%	8%	0.08	11%	13%	-0.10
ELD program participant (level 4 or 5)	18%	19%	-0.03	15%	16%	-0.06
Attendance rate (ninth grade)	92%	92%	0.00	85%	84%	0.04
Average GPA (ninth grade)	1.51	1.58	-0.10	1.37	1.34	0.03
Average SB Grade 8 z-score: ELA	-0.44	-0.35	-0.12	-0.46	-0.47	0.01
Average SB Grade 8 z-score: math	-0.52	-0.42	-0.14	-0.44	-0.38	-0.08

Note. ELA = English language arts; ELD = English language development; GPA = grade point average; SB = Smarter Balanced; SMD = standardized mean difference.

The SB scale score was standardized on the basis of the districtwide mean and standard deviation. The SMD was calculated using the Cox index for dichotomous measures and Hedge’s *g* for continuous measures.

Table 2. Comparison of the Total Student Sample to the Sample of Test Takers for Algebra 1, by Treatment Group

Student characteristics	Online classes		Teacher-directed classes		SMD between online and teacher-directed test takers
	Total student sample	Test takers	Total student sample	Test takers	
Number of students	305	181	308	222	
Female	47%	43%	43%	42%	0.02
Ethnicity: African American/Black	10%	11%	12%	11%	0.01
Ethnicity: Latinx/Hispanic	81%	78%	80%	81%	-0.10
Ethnicity: other	9%	11%	8%	8%	0.17
National school lunch–eligible	80%	76%	80%	81%	-0.18
Gifted/talented program	6%	7%	8%	9%	-0.17
Student with a disability	10%	7%	8%	8%	-0.04
ELD program participant (level 4 or 5)	18%	19%	19%	19%	0.00
Attendance rate (ninth grade)	92%	93%	92%	93%	0.08
Average GPA (ninth grade)	1.51	1.60	1.58	1.64	-0.06
Average SB Grade 8 z-score: ELA	-0.44	-0.37	-0.35	-0.31	-0.07
Average SB Grade 8 z-score: math	-0.52	-0.48	-0.42	-0.41	-0.09

Note. ELA = English language arts; ELD = English language development; GPA = grade point average; SB = Smarter Balanced; SMD = standardized mean difference.

The SB scale score was standardized on the basis of the districtwide mean and standard deviation. The SMD was calculated using the Cox index for dichotomous measures and Hedge’s *g* for continuous measures.

Table 3. Comparison of the Total Student Sample to the Sample of Test Takers for English 9, by Treatment Group

Student characteristics	Online classes		Teacher-directed classes		SMD between online and teacher-directed test takers
	Total student sample	Test takers	Total student sample	Test takers	
Number of students	564	324	560	336	
Female	35%	37%	33%	32%	0.15
Ethnicity: African American/Black	8%	8%	9%	8%	0.00
Ethnicity: Latinx/Hispanic	85%	85%	83%	84%	0.02
Ethnicity: other	7%	7%	9%	8%	-0.03
National school lunch–eligible	89%	90%	90%	88%	0.07
Gifted/talented program	12%	15%	12%	12%	0.12
Student with a disability	11%	9%	13%	13%	-0.23
ELD program participant (level 4 or 5)	15%	14%	16%	15%	-0.08
Attendance rate (ninth grade)	85%	88%	84%	87%	0.04
Average GPA (ninth grade)	1.37	1.43	1.34	1.47	-0.06
Average SB Grade 8 z-score: ELA	-0.46	-0.36	-0.47	-0.46	0.12
Average SB Grade 8 z-score: math	-0.44	-0.35	-0.38	-0.35	0.01

Note. ELA = English language arts; ELD = English language development; GPA = grade point average; SB = Smarter Balanced; SMD = standardized mean difference.

The SB scale score was standardized on the basis of the districtwide mean and standard deviation. The SMD was calculated using the Cox index for dichotomous measures and Hedge’s *g* for continuous measures.

Data and Measures

We collected the following primary and extant data to address the research questions:

- Weekly teacher logs to document modes of instructional delivery and content coverage;
- An end-of-course teacher survey to measure teacher perceptions of instructional features aligned with the theory of action (see Appendix 1);
- An end-of-course student survey to measure student instructional experiences aligned with the theory of action;
- A study-developed end-of-course test to measure student knowledge of course content; and
- District extant data that included student background characteristics, eighth-grade state test scores, ninth-grade academic performance, and the final grade students received in the credit recovery course.

Teacher logs. The logs were administered online at the end of each week of the 5-week summer school sessions. We received at least one log for every class. Across all five logs per class, teachers completed 93% of the Algebra 1 logs (91% for online classes and 94% for teacher-directed classes) and more than 97% of the English 9 logs (97% for the online and teacher-directed classes). Using a 5-point scale,⁶ the logs asked teachers to report how often they engaged in different instructional activities, such as large-group instruction, small-group instruction, and individual tutoring. Another set of questions asked how often most students in the class engaged in different modes of instruction. In addition, using a 4-point scale,⁷ the logs asked teachers to report on how many students in the class covered specific types of content during the week.

To estimate the cost of the credit recovery classes, we gathered information from the teacher logs, where teachers reported how much time they spent on credit recovery class-related activities, both in class and outside of class.

Teacher survey. The teacher survey was administered online at the end of the summer term. There was only one session of credit recovery classes per summer. The response rate was 100% for Algebra 1 and 97% for English 9 (97% for both online and teacher-directed classes). The surveys included a series of statements about instruction and asked teachers to report how

⁶ The 5-point scale used in the logs had the following response options: never (0%), a little (1%–25%), sometimes (26%–50%), often (51%–75%), and a lot (76%–100%).

⁷ The 4-point scale used in the logs had the following response options: no students, few students, some students, and most students.

much they agreed with each statement, on a 4-point scale,⁸ for their credit recovery class (see Table 4). Most of the survey items were adapted from items used in a study of personalized learning (Pane et al., 2015).

With our final sample, which included the summer and school-year study classes,⁹ we conducted exploratory and confirmatory factor analyses to develop survey measures for four instructional features aligned with the theory of action:

- *Individualized pacing* measures the extent to which students can work through the course content at their own pace.
- *Connection of content to student needs* measures the extent to which the course content and materials are aligned to individual student learning needs and experiences.
- *Instructional support* measures the extent to which students can get individualized support when they need it.
- *Performance feedback* measures the extent to which students know the learning goals of class assignments and activities and receive timely feedback on their class performance.

In addition to these four instructional features, we measured the extent to which teachers felt prepared and supported to teach a specific course. For each of the teacher survey constructs, we generated factor scores that were standardized based on the total sample mean and standard deviation. The standardization included Algebra 1 and English 9 classes in the summer and school-year samples. All formal statistical tests of group differences (described in the next section) are based on the standardized scale scores.

For presentation purposes in the briefs, we converted the group mean standardized scale scores into an index score that ranges from 0 to 100, where an index score of 50 represents the mean score across all credit recovery classes in the study. The index approximates a percentile rank based on a normal distribution and is similar to the improvement index used by the What Works Clearinghouse (2020). For example, a group mean index score of 60 indicates that the group mean is 10 percentile points above the overall mean.

We also asked teachers about the grading criteria they used to determine final grades for the credit recovery class so that we could compare grading standards in the online and teacher-directed classes. We grouped teacher responses into three categories: behavior-related criteria

⁸ The 4-point scale used in the surveys had the following response options: strongly disagree, disagree, agree, and strongly agree.

⁹ Teachers from the school-year sample were included in our assessment of the survey item measurement properties and calculation of factor scores so that the survey measures could be compared across the summer and school-year samples if desired. The results presented in the briefs and this technical supplement are based on the summer school classes only.

(e.g., attendance and effort), class assignments (e.g., classwork, homework, and essays), and test/quiz performance.

To gain complete information for the cost analysis, we asked teachers about the time they allocated for professional development activities related to the credit recovery courses, time spent by other school staff or volunteers helping with the courses, and the equipment and materials used for the courses (e.g., computers, textbooks).

Table 4. Teacher Survey Questions for Each Instructional Feature

Individualized pacing. Please indicate your level of agreement with the following statements about the curriculum and instruction in this class. (Internal consistency = 0.91)
Students must show that they understand a topic before they can move to a new topic.
Different students work on different topics or skills at the same time.
Students can work through instructional material at a faster or slower pace than other students in this class.
Students have opportunities to review or practice new material until they fully understand it.
Students can access instructional materials both in and outside the classroom.
Students keep track of their own learning progress.
Connection of content to student needs. Please indicate your level of agreement with the following statements about the instructional support available to students in this class. (Internal consistency = 0.94)
Various materials or instructional approaches are available to accommodate individual student needs or interests.
Course content is adapted to meet students’ needs by providing additional assignments, resources, and activities for remediation or enrichment.
The course content connects what students are learning with experiences they have throughout the rest of the school day or outside school.
Instructional support. Please indicate your level of agreement with the following statements about the instructional support available to students in this class. (Internal consistency = 0.97)
A teacher is available to provide individual tutoring to students during class.
A teacher is available to provide coaching or support to students while they are working together in groups or individually.
If students have trouble understanding material, they can get help quickly.
Performance feedback. Please indicate your level of agreement with the following statements about students in this class. (Internal consistency = 0.97)
Students receive immediate feedback on problem solutions.
Students receive feedback about their strengths and weaknesses in the course.
When students work on an assignment or activity, they know what the goals of the assignment or activity are.

Teacher preparation/support. Please indicate your level of agreement with the following statements about teaching this class. (Internal consistency = 0.88)

I felt well prepared to teach this class.
I had the necessary materials (e.g., textbooks, supplies) to teach this class.
I had the necessary support from peers or leaders to teach this class.
I had access to high-quality assessment data that helped me adapt the pace or content of instruction to meet students' needs.
I had the necessary skills and experience to use data to guide my instruction.

Note. All items had the following response options: strongly disagree, disagree, agree, or strongly agree. Cronbach's alpha was used to calculate internal consistency values.

Student survey. The student survey was administered in class by the study team during the last week of the summer term. The response rate was 66% for Algebra 1 (60% for online classes and 73% for teacher-directed classes) and 59% for English 9 (57% for online classes and 61% for teacher-directed classes).¹⁰ The survey included a series of statements about the class and asked students to report how much they agreed with each statement, on a 4-point scale,¹¹ for their credit recovery class (see Table 5). Most of the survey items were adapted from items used in a study of student engagement (Skinner et al., 2009) or an earlier credit recovery study that took place in Chicago (Heppen et al., 2016). With our final sample, including the summer and school-year study classes,¹² we conducted exploratory and confirmatory factor analyses to develop survey measures for five student experiences aligned with the theory of action:

- *Behavioral engagement* measures the extent to which students made efforts and took actions to learn in class.
- *Emotional engagement* measures the extent to which students felt enthusiasm and enjoyment in the class.
- *Personalized instruction* measures the extent to which students thought their teacher provided them with additional instruction when they needed help.

¹⁰ The lower than desired response rate partially reflects the fact that the credit recovery classes were poorly attended; on any given day a significant percentage of the students were not in class. In some cases, certain students rarely attended. For students absent on the day we administered the survey, we provided teachers with copies of the survey and ask them to give them to the students when they did attend, and then mail the completed surveys to us. This resulted in only a small number of additional surveys.

¹¹ The 4-point scale had the following response options: strongly disagree, disagree, agree, and strongly agree.

¹² Students from the school-year sample were included in our assessment of the survey item measurement properties and calculation of factor scores so that the survey measures could be compared across the summer and school-year samples if desired. The results presented in the briefs and this technical supplement are based on the summer school classes only.

- *Academic challenge* measures the extent to which students thought the class was challenging.
- *Clarity of class expectations* measures the extent to which students thought the teacher had high expectations and that they understood how the class work aligned with what they should be learning.

For each of the student survey constructs, we generated factor scores that were standardized based on the total sample mean and standard deviation, including classes from Algebra 1 and English 9 in the summer and school-year samples. All formal statistical tests of group differences among the summer school classes are based on the standardized scale scores.

As with the teacher surveys, we converted the group mean standardized scale scores into an index score that ranges from 0 to 100, where an index score of 50 represents the mean score across all credit recovery classes in the study. The index approximates a percentile rank based on a normal distribution and is similar to the improvement index used by the What Works Clearinghouse (2020). A group mean index score of 60, for example, indicates that the group mean is 10 percentile points above the overall mean.

Table 5. Student Survey Questions for Each Instructional Experience

Behavioral engagement. How much do you agree with the following statements about your time in this class? (Internal consistency = 0.84)
I try hard to do well in this class.
In class, I work as hard as I can.
When I’m in class, I participate in class discussions.
I pay attention in class.
When I’m in class, I listen very carefully.
Emotional engagement. How much do you agree with the following statements about your time in this class? (Internal consistency = 0.83)
When I’m in class, I feel good.
When we work on something in class, I feel interested.
Class is fun.
I enjoy learning new things in class.
When we work on something in class, I get involved.
Personalized instruction. How much do you agree with the following statements about your teacher for this class? (Internal consistency = 0.90)
My teacher helped me catch up if I was behind.
My teacher was willing to give extra help on work if I needed it.
My teacher noticed whether I had trouble learning something.

My teacher gave me specific suggestions about how I could improve my work in this class.
My teacher explained things in a different way if I did not understand something in class.
Academic challenge. How much do you agree with the following statements about this class? (Internal consistency = 0.68)
I found the work challenging.
The class really made me think.
I had to work hard to do well in this class.
The assignments often required me to explain my answers.
Clarity of class expectations. How much do you agree with the following statements about this class? (Internal consistency = 0.87)
My teacher expected everyone to work hard.
My teacher expected us to become better thinkers, not just memorize things.
I learned a lot from feedback on my work in class.
It was clear what I needed to do to get a good grade in this class.
The work that we did was good preparation for the tests in this class.
The class assignments helped me learn the course material.
I knew what my teacher wanted me to learn in this class.

Note. All items had the following response options: strongly disagree, disagree, agree, or strongly agree. Cronbach’s alpha was used to calculate internal consistency values.

Student test. The student test was administered in class by the study team during the last week of the summer term. The response rate was 66% for Algebra 1 (59% for online classes and 72% for teacher-directed classes) and 59% for English 9 (57% for online classes and 60% for teacher-directed classes). The Algebra 1 test included 20 multiple choice items taken from the pool of publicly released Grade 8 and Grade 12 National Assessment of Educational Progress (NAEP) mathematics assessments. We selected items that covered the range of topics typically taught in a first-year algebra course, including some items that cover important pre-algebra content. The English 9 test included 20 multiple choice items taken from a pool of publicly released items from Ohio’s Grade 8 and high school English language arts state assessment.¹³ The test included items about two literary texts and two informational texts. For the Algebra 1 and English 9 tests, we used a two-parameter item response model to create student scale scores. We standardized the scale scores based on the total sample mean and standard deviation

¹³ We used items from the Ohio assessments because the assessment had a good selection of publicly available reading comprehension items and an AIR English language arts content expert confirmed that the items covered California content standards for Grade 9 and 10 English language arts. Both California and Ohio adopted the Common Core State Standards.

(separately for each subject).¹⁴ All formal statistical tests of group differences (described in the next section) are based on the standardized scale scores.

As before, for presentation purposes in the briefs, we converted the group mean standardized scale scores into an index score that ranges from 0 to 100, where an index score of 50 represents the mean score across all credit recovery classes in the study. The index approximates a percentile rank based on a normal distribution and is similar to the improvement index used by the What Works Clearinghouse (2020).

Final course grade. The district provided us with the final course grades that students received in their credit recovery classes. Per district policy, we defined any student with a grade of D or better as having passed the class and recovered the course credit. Some students who were enrolled in a study class at the start of the summer term did not receive a final grade because they either dropped the class or took an incomplete. We coded all students without a final grade as having an “incomplete” and as not passing the class during the summer.

District extant data. For all students in the study, the district provided us with data on student characteristics, eighth-grade state assessment English language arts and math scores, course grades in ninth grade, and school attendance in ninth grade. We used this information to check baseline equivalence (see Table 1) and as covariates in the student outcome models (see description of analysis in the next section). For course grades in ninth grade, we calculated each student’s grade point average (GPA) as the average GPA in the fall and spring semesters.

Analysis Approach

We conducted separate analyses for Algebra 1 and English 9. To compare the instructional features of the online and teacher-directed classes, we used a linear regression model with school fixed effects (and the cohort for English 9) to estimate the average within-school difference between online and teacher-directed classes. For student instructional experience and outcomes, we analyzed the data based on the type of class to which students were randomly assigned to estimate the intent-to-treat (ITT) average effect.¹⁵ To estimate the average effect, we used a linear regression model for the test score outcome and a logistic regression model for whether students passed or did not pass the class. The regression models

¹⁴ For the summer terms, the Algebra 1 test score had an empirical marginal reliability of 0.66 (internal consistency = 0.59) and the English 9 test score had an empirical marginal reliability of 0.77 (internal consistency = 0.75).

¹⁵ Approximately 90% of the students were enrolled in the class to which they were assigned.

controlled for all student background characteristics¹⁶ listed in Table 1 and included fixed effects for the randomization blocks, which account for school (and cohort for English 9).

Analytic Models

For analyses of class instructional features, we estimated the following fixed-effects linear regression model for class i in school j :

$$(1) \quad Y_i = \beta_0 + \beta_1 T_i + \mathbf{S}'_j \boldsymbol{\gamma}_j + e_i,$$

where T = dichotomous indicator for online (1) or teacher-directed (0) class

and \mathbf{S} = vector of term-by-school fixed effects

For analyses of the student instructional experience measures, student content knowledge measures, and total credit accumulation measures (the continuous outcome measures), we estimated the following fixed-effects linear regression model for student i in random assignment block j :

$$(2) \quad Y_i = \beta_0 + \beta_1 T_i + \mathbf{X}'_i \boldsymbol{\beta}_x + \mathbf{B}'_j \boldsymbol{\gamma}_j + e_i,$$

where T = dichotomous treatment assignment indicator

and \mathbf{X} = vector of the student characteristics listed in the baseline equivalence tables

and \mathbf{B} = vector of block fixed effects, which include school and cohort

For analyses of credit recovery (the binary outcome measure), we estimated logistic regression models that parallel Equation 2 with a logit-link function.

In addition to the main analyses, we examined the extent to which the overall average effects differed across student characteristics (moderator analysis) by estimating separate models for each moderator, where an interaction term between treatment and the moderator was added to Equation 2. In particular, we tested for differences based on the following student characteristics:

- Ninth-grade attendance rate (school year prior to the summer course)
- Ninth-grade GPA (school year prior to the summer course)
- Eighth-grade state assessment scores for math and ELA.

¹⁶ Due to potential non-linearities in the attendance rate, we included dichotomous indicators for the following attendance rate categories in the model instead of the actual attendance rate: less than 75%, 75%–84%, 85%–89%, 90–94%, and 95%–100%.

Missing Data Approach

Missing data are a potential concern for the eighth-grade state assessment scores (covariates) and the student survey and test (outcomes). Response rates for the student survey and test administered toward the end of the summer term are reported in Table 6. To account for potential bias due to missing data, we used multiple imputation with chained equations (M = 20) and predictive mean matching. We ran separate imputation models by subject and treatment condition. The imputation models included all covariates, randomization blocks, and student outcomes. Fortunately, the sample of students with test and survey data have similar background characteristics to the full student sample (see Tables 2 and 3), which gives us some confidence that the outcome data are missing at random. To examine whether our results are sensitive to the imputation of missing values for the student survey and test outcomes, we present results from analyses that exclude missing data in Tables 21 and 22 in the sensitivity analysis section.

Table 6. Response Rates for the Student Survey and Test

Outcome	Total sample	Treatment group	Control group	Treatment – control difference
Algebra 1				
Student survey	66%	60%	73%	-13%
Student test	66%	59%	72%	-13%
English 9				
Student survey	59%	57%	61%	-4%
Student test	59%	57%	60%	-3%

Impact Results

Main Impact Estimates

Table 7. Estimated Effect on Instructional Features as Measured by the Teacher Survey

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra 1					
Individualized pacing	0.837	-0.206	1.043	0.290	0.003
Performance feedback	0.599	0.188	0.411	0.159	0.021
Instructional support	0.485	0.096	0.389	0.228	0.110
Connection of content to student needs	0.243	0.219	0.025	0.199	0.903
English 9					
Individualized pacing	0.186	-0.202	0.388	0.219	0.084
Performance feedback	-0.310	0.253	-0.563	0.195	0.006
Instructional support	-0.182	0.061	-0.244	0.199	0.227
Connection of content to student needs	-0.402	0.408	-0.810	0.203	0.000

Table 8. Estimated Effect on Student Experiences as Measured by the Student Survey

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra 1					
Behavioral engagement	-0.014	0.080	-0.093	0.096	0.334
Emotional engagement	-0.136	0.024	-0.159	0.096	0.099
Personalized instruction	0.040	-0.056	0.096	0.086	0.266
Academic challenge	0.106	0.162	-0.056	0.093	0.553
Clarity of class expectations	-0.053	0.071	-0.124	0.093	0.184
English 9					
Behavioral engagement	0.160	-0.018	0.178	0.085	0.041
Emotional engagement	0.140	-0.040	0.180	0.088	0.045
Personalized instruction	0.137	-0.110	0.247	0.082	0.004
Academic challenge	0.049	-0.049	0.099	0.082	0.232
Clarity of class expectations	0.083	-0.035	0.118	0.081	0.150

Table 9. Estimated Effect on Student Test Scores

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra test score	0.130	-0.050	0.181	0.101	0.078
English test score	-0.057	-0.052	-0.005	0.084	0.952

Table 10. Estimated Effect on Student Credit Recovery Rate

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect (logit)	Standard error (logit)	p-value
Algebra course pass rate	62.3%	68.0%	-0.252	0.213	0.236
English course pass rate	52.3%	74.0%	-0.954	0.155	0.000

Results From Moderator Analysis

Table 11. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Experiences

Outcome	Estimated effect	Standard error	p-value
Algebra 1			
Behavioral engagement	-0.04	0.195	0.836
Emotional engagement	-0.063	0.187	0.737
Personalized instruction	0.025	0.183	0.893
Academic challenge	-0.03	0.195	0.876
Clarity of class expectations	-0.081	0.183	0.656
English 9			
Behavioral engagement	0.139	0.165	0.400
Emotional engagement	0.201	0.175	0.253
Personalized instruction	0.062	0.177	0.726
Academic challenge	0.149	0.177	0.400
Clarity of class expectations	0.085	0.178	0.631

Note. For the attendance rate, we included dichotomous indicators for the following attendance rate categories in the model, along with interactions with the treatment indicator, instead of the actual attendance rate: less than 75%, 75%–84%, 85%–89%, 90%–94%, and 95%–100%. The attendance moderator effect presented in the table is for the maximal contrast between the 95%–100% attendance rate group and the <75% attendance group.

Table 12. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Test Scores

Outcome	Estimated effect	Standard error	p-value
Algebra test score	0.071	0.184	0.701
English test score	-0.284	0.145	0.052

Note. For the attendance rate, we included dichotomous indicators for the following attendance rate categories in the model, along with interactions with the treatment indicator, instead of the actual attendance rate: less than 75%, 75%–84%, 85%–89%, 90%–94%, and 95%–100%. The attendance moderator effect presented in the table is for the maximal contrast between the 95%–100% attendance rate group and the <75% attendance group.

Table 13. Estimated Interaction Effect of Treatment and Ninth-Grade Attendance on Student Credit Recovery Rate

Outcome	Estimated effect (logit)	Standard error	p-value
Algebra course pass rate	0.344	0.457	0.452
English course pass rate	-0.438	0.382	0.252

Note. For the attendance rate, we included dichotomous indicators for the following attendance rate categories in the model, along with interactions with the treatment indicator, instead of the actual attendance rate: less than 75%, 75%–84%, 85%–89%, 90%–94%, and 95%–100%. The attendance moderator effect presented in the table is for the maximal contrast between the 95%–100% attendance rate group and the <75% attendance group.

Table 14. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Experiences

Outcome	Estimated effect	Standard error	p-value
Algebra 1			
Behavioral engagement	-0.132	0.171	0.443
Emotional engagement	-0.205	0.162	0.209
Personalized instruction	0.03	0.149	0.843
Academic challenge	0.071	0.153	0.643
Clarity of class expectations	-0.05	0.154	0.747
English 9			
Behavioral engagement	0.053	0.104	0.612
Emotional engagement	0.075	0.106	0.482
Personalized instruction	0.018	0.118	0.882
Academic challenge	0.201	0.111	0.075

Clarity of class expectations	0.104	0.108	0.337
-------------------------------	-------	-------	-------

Table 15. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Test Scores

Outcome	Estimated effect	Standard error	p-value
Algebra test score	0.254	0.131	0.057
English test score	0.057	0.097	0.556

Table 16. Estimated Interaction Effect of Treatment and Ninth-Grade GPA on Student Credit Recovery Rate

Outcome	Estimated effect (logit)	Standard error	p-value
Algebra course pass rate	0.302	0.334	0.366
English course pass rate	-0.629	0.233	0.007

Table 17. Estimated Interaction Effect of Treatment and Eighth-Grade English Assessment on Student Experiences (English 9 Classes Only)

Outcome	Estimated effect	Standard error	p-value
Behavioral engagement	0.044	0.091	0.626
Emotional engagement	0.111	0.095	0.246
Personalized instruction	0.001	0.094	0.997
Academic challenge	0.055	0.092	0.549
Clarity of class expectations	0.046	0.093	0.62

Table 18. Estimated Interaction Effect of Treatment and Eighth-Grade English Assessment on Student Test Scores and Student Credit Recovery Pass Rate (English 9 Classes Only)

Outcome	Estimated effect	Standard error	p-value
English test score	-0.004	0.096	0.964
English course pass rate	0.150	0.199	0.451

Table 19. Estimated Interaction Effect of Treatment and Eighth-Grade Math Assessment on Student Experiences (Algebra 1 Classes Only)

Outcome	Estimated effect	Standard error	p value
Behavioral engagement	-0.333	0.148	0.028
Emotional engagement	-0.419	0.132	0.002
Personalized instruction	-0.071	0.133	0.594
Academic challenge	-0.014	0.139	0.917
Clarity of class expectations	-0.25	0.127	0.050

Table 20. Estimated Interaction Effect of Treatment and Eighth-Grade Math Assessment on Student Test Scores and Student Credit Recovery Pass Rate (Algebra 1 Classes Only)

Outcome	Estimated effect	Standard error	p value
Algebra test score	0.036	0.11	0.743
Algebra course pass rate	-0.755	0.349	0.031

Results From Sensitivity Analysis

This section presents results from the two types of sensitivity analyses that we conducted. First, to demonstrate that the multiple imputations of missing student survey and test measures do not substantively affect our results, we present results from an analysis of student surveys and test measures that excludes cases with missing data in Tables 21 and 22. Compared with the main results presented in Tables 8 and 9, two findings differ substantively when cases with missing data are excluded: (1) the statistically significant positive effect on the two engagement measures for English 9 are no longer significant, and (2) the estimated effect on the test score for Algebra 1 increases and becomes statistically significant.

Second, to demonstrate that our results are not substantively affected by the fact that some schools did not use the study-provided lottery lists, we examined how average effect estimates differed based on whether a school did or did not use the study-provided lottery lists. We examined this in two ways. First, we estimated separate models on subsamples defined by whether the school used our list. Second, we estimated a model on the full sample including an interaction term between treatment and a dichotomous indicator for whether the school did or did not use the study-provided lottery list. No significant interactions were observed. Tables 25–27 present estimates for the subsample of schools that used the study-provided randomization list. For the schools that participated in the summer of 2018 (English courses only), 13 schools (675 students) used the study-provided list and two schools (55 students) did not. For the summer 2019 Algebra courses, 11 schools (501 students) used the study-provided

list and two schools (112 students) did not. For the summer 2019 English courses, eight schools (316 students) used the study-provided list and two schools (78 students) did not.

Table 21. Estimated Effect on Student Experiences as Measured by the Student Survey, Without Imputation of Missing Data

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra 1					
Behavioral engagement	-0.019	0.026	-0.045	0.107	0.675
Emotional engagement	-0.130	0.034	-0.165	0.112	0.143
Personalized instruction	0.023	-0.005	0.028	0.106	0.793
Academic challenge	0.123	0.217	-0.094	0.105	0.372
Clarity of class expectations	-0.063	0.096	-0.159	0.111	0.153
English 9					
Behavioral engagement	0.094	0.071	0.022	0.085	0.791
Emotional engagement	0.069	-0.019	0.088	0.089	0.325
Personalized instruction	0.102	-0.170	0.271	0.087	0.002
Academic challenge	-0.031	-0.060	0.028	0.091	0.755
Clarity of class expectations	0.053	-0.083	0.136	0.089	0.125

Table 22. Estimated Effect on Student Test Scores, Without Imputation of Missing Data

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra test score	0.177	-0.048	0.225	0.102	0.028
English test score	-0.058	0.016	-0.074	0.075	0.325

Table 23. Sensitivity Analysis Estimated Effect on Student Experiences as Measured by the Student Survey for Schools That Complied With Randomization Procedure

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra 1					
Behavioral engagement	-0.035	0.084	-0.119	0.105	0.260
Emotional engagement	-0.193	-0.002	-0.191	0.107	0.078
Personalized instruction	-0.028	-0.138	0.110	0.104	0.292
Academic challenge	0.081	0.175	-0.095	0.101	0.351
Clarity of class expectations	-0.116	0.042	-0.159	0.106	0.139
English 9					
Behavioral engagement	0.176	-0.022	0.198	0.092	0.035
Emotional engagement	0.129	-0.051	0.180	0.095	0.066
Personalized instruction	0.103	-0.098	0.201	0.082	0.017
Academic challenge	0.046	-0.042	0.088	0.084	0.301
Clarity of class expectations	0.074	-0.024	0.097	0.087	0.267

Table 24. Sensitivity Analysis Estimated Effect on Student Test Scores for Schools That Complied With Randomization Procedure

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect	Standard error	p-value
Algebra test score	0.084	-0.117	0.201	0.115	0.087
English test score	-0.096	-0.112	0.015	0.093	0.869

Table 25. Sensitivity Analysis Estimated Effect on Student Credit Recovery Rate for Schools That Complied With Randomization Procedure

Outcome	Treatment mean	Control mean (adjusted)	Estimated effect (logit)	Standard error (logit)	p-value
Algebra course pass rate	61.0%	66.1%	-0.220	0.243	0.366
English course pass rate	51.6%	73.4%	-0.949	0.163	0.000

Cost Analysis

For the cost analysis, we examined the types and quantities of personnel and non-personnel resources needed to deliver the Edgenuity online credit recovery classes to students compared with typical, teacher-directed credit recovery classes, as well as the corresponding costs of those resources. We used the ingredients approach to costing out educational services, as developed by Levin and colleagues (2018). This approach involves identifying the comprehensive list of “ingredients”—personnel and non-personnel resources such as instructor time, computers, and textbooks—associated with providing credit recovery courses, including their quantities and unit prices. Quantities of ingredients and unit prices were used to cost out individual ingredients, which were then aggregated to provide an estimate of the cost of the credit recovery classes. We focused on the ingredients that were likely to differ between the two types of courses (such as instructional technology and time spent by teachers on various activities), and omitted some ingredients that we assumed would be the same between the two types of courses (such as classroom space).

Cost Methodology

Data used to develop cost estimates came from several sources. Information on teacher time came from teacher logs, in which teachers were asked to report how much time they spent during the past week in class and on additional, class-related activities outside of class time, such as planning lessons, developing course materials, and grading. In addition, the team administered teacher surveys that included items asking about time spent on professional development related to the courses, time spent by other school staff or volunteers helping with the courses, and non-personnel equipment and materials used for the courses (such as computers, presentation equipment, and textbooks). Finally, we conducted an interview with school administrators to assess any differences in administrative time required for setting up and managing the two courses.

We estimated the cost per student for all credit recovery classes offered as part of this study and came up with an average cost per student across all online classes and an average cost per student across all teacher-directed classes.¹⁷ We calculated the cost per student using a constant class size of 25 students. The results presented in Brief 4 assume a constant class size

¹⁷ Averages were calculated as conditional averages using a regression model containing school-by-term fixed effects. These school-by-term fixed effects control for unobserved differences in cost across schools and terms. Essentially, the school-by-term fixed effects compare the cost of classes within school and term. Separate regressions were run for summer and school-year courses and for Algebra and English courses.

of 25 students, which provides a better comparison of the costs one might expect when implementing the courses outside of the study context.¹⁸

For cost calculations and reporting, we classified costs in two ways: fixed versus variable costs and district-incurred versus teacher-incurred costs. We discuss these classifications in this section. Table 26 provides the classification decision for each of the main resources included in the cost analysis, along with the estimated per-unit cost for each resource.

Classifying Fixed and Variable Costs

To make cost estimates based on a constant class size, we had to make assumptions about which categories of cost were fixed and which were variable. Fixed costs are costs that remain the same regardless of the number of students in the class. For fixed costs, the cost per student decreases as the number of students increases. In contrast, variable costs change with the number of students. If each student in the class needs a textbook or computer, as the number of students increases, the cost of textbooks and computers increases. Because the costs increase or decrease proportionally to the number of students, the per-student cost for variable costs does not change with class size.

We categorized some costs as semi-variable. These are costs that change with class size, but the per-unit cost decreases somewhat with additional students. For example, we classified grading of student work as a semi-variable cost under the assumption that it takes longer to grade the first assignment than it does each additional assignment. A teacher might have to construct a rubric or a key for grading, which would be a fixed cost, but then scoring assignments would be a variable cost.

After classifying costs as fixed, variable, or semi-variable, the AIR team calculated per-student costs using a constant class size of 25. For fixed costs, this simply meant dividing the total cost by 25 rather than by the actual class size. For variable costs, this meant retaining the same per-pupil cost as calculated using the actual class size. For semi-variable costs, we estimated the per-pupil costs if the resource had been classified as fixed and then estimated per-pupil cost if the resource was classified as variable. Finally, we took the average of both per-pupil costs.

¹⁸ To participate in this study, schools were required to offer two sections of the same course, such that one section could be offered as a teacher-directed course and the other course could be offered as an online course. In some instances, this requirement created class sizes that were smaller than they would have been if schools had not been required to create two course sections. In addition to smaller-than-typical class sizes, the online courses ended up having slightly smaller class sizes than the teacher-directed courses, on average, during the summer sessions. Because the class size affects the calculated cost per student, in addition to estimating costs per student using actual class sizes, we predicted costs per student using a constant class size of 25 students. Our preferred results are those in which a constant class size of 25 students was used. When using a constant, predicted, per-pupil class size, averages weighted by the constant class size are the same as unweighted averages.

Classifying District- Versus Teacher-Incurred Costs

In addition to looking at overall costs, we examined costs based on whether the costs were paid directly by the district (district-incurred costs) or whether the costs would not affect the district budget. In the case of this analysis, the costs for which the district would not have to pay are in the form of extra time that teachers provide outside of their contracted hours (teacher-incurred costs). For example, if a teacher spends time helping students outside of class time, that is a resource that contributes to the success of the intervention, but the teacher does not get paid more for this additional time. The same is true for the amount of time that teachers spend creating lesson plans, grading student assignments, developing course materials, and other activities, such as communicating with parents or collaborating with other teachers. These teacher-incurred costs are opportunity costs to the teacher but they do not directly affect district expenditures.

To distinguish between teachers’ contracted hours and the additional time teachers spent for the credit recovery courses, we assumed that for every 6 hours of class time, teachers get 1 hour of paid planning time covered by their contract. For a summer class, this amounts to approximately 10 hours of contracted time per class for teachers to spend on out-of-class activities. Hours spent on out-of-class activities beyond the allotted contractual planning time were considered teacher-incurred costs.

Table 26. Description of Resources Accounted for in the Cost Analysis and Their Categorization

Resource description	Data source	District or teacher cost	Fixed or variable cost	Cost per unit
Personnel resources				
Teacher class time	Teacher log	District	Fixed	\$60.10 per hour
Teacher lesson plans	Teacher log	Teacher	Fixed	\$60.10 per hour
Teacher grading	Teacher log	Teacher	Semi-variable	\$60.10 per hour
Teacher materials development	Teacher log	Teacher	Fixed	\$60.10 per hour
Teacher-provided additional assistance to students	Teacher log	Teacher	Semi-variable	\$60.10 per hour
Teacher other activity	Teacher log	Teacher	Semi-variable	\$60.10 per hour
Teacher professional development	Teacher survey	District	Fixed	\$60.10 per hour
Teaching assistant classroom assistance	Teacher survey	District	Fixed	\$21.70 per hour

Resource description	Data source	District or teacher cost	Fixed or variable cost	Cost per unit
Special education teacher classroom assistance	Teacher survey	District	Fixed	\$61.06 per hour
English learner teacher classroom assistance	Teacher survey	District	Fixed	\$60.10 per hour
Principal or assistant principal classroom assistance	Teacher survey	District	Fixed	\$126.44 per hour
Principal or assistant principal course setup	Administration interview	District	Fixed	\$126.44 per hour
Other administrative staff classroom assistance	Teacher survey	District	Fixed	\$38.13 per hour
Other administrative staff course setup	Administration interview	District	Fixed	\$38.13 per hour
Information technology (IT) support course setup	Administration interview	District	Fixed	\$51.06 per hour
Other teacher or instructional coach classroom assistance	Teacher survey	District	Fixed	\$60.10 per hour
Non-personnel resources				
Computers	Teacher survey	District	Variable	\$19.69 per student
Electronic whiteboard	Teacher survey	District	Fixed	\$100 per class
Digital projector	Teacher survey	District	Fixed	\$15 per class
Dry-erase whiteboard	Teacher survey	District	Fixed	\$5 per class
Chalkboard	Teacher survey	District	Fixed	\$3.75 per class
Digital overhead projector	Teacher survey	District	Fixed	\$9.38 per class
Textbooks	Teacher survey	District	Variable	\$8.33 per student
Other published materials	Teacher survey	District	Variable	\$0.50 per student
Edgenuity license	Study design	District	Variable	\$34.38 per student
Edgenuity training and support	Study design	District	Fixed	\$275 per class

Note. *Fixed costs* do not change with class size (e.g., the cost of a digital projector is constant regardless of whether there are 10 or 20 students in the class). *Variable costs* change proportionately with class size (e.g., each student gets a textbook, so the number and cost of textbooks increases proportionately with the number of students). *Semi-variable costs* are somewhere in between fixed and variable costs. Costs increase with class size, but not proportionately.

References

- Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness*, *10*(2), 272–296. <https://doi.org/10.1080/19345747.2016.1168500>
- Levin, H., McEwan, P., Belfield, C., Bowden, B. & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. (3rd ed.). SAGE Publications, Inc.
- Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). *Continued progress: Promising evidence on personalized learning*. RAND Corporation. http://www.rand.org/content/dam/rand/pubs/research_reports/RR1300/RR1365/RAND_RR1365.pdf
- Rickles, J., Heppen, J. B., Allensworth, E., Sorensen, N., & Walters, K. (2018). Online credit recovery and the path to on-time high school graduation. *Educational Researcher*, *47*(8), 481–491.
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children’s behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493–525. <https://doi.org/10.1177/0013164408323233>
- What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook, version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. This report is available on the What Works Clearinghouse website at <https://ies.ed.gov/ncee/wwc/handbooks>.

Appendix 1. Theory of Action

Theory of Action

The intervention for this study was an Algebra 1 or English 9 (first or second semester) online curriculum for the credit recovery course,¹⁹ where an online provider supplied the main course content and curriculum, and the school provided the appropriate, credentialed in-class teacher who could supplement the digital instruction. For both the intervention and business-as-usual (BAU) conditions, students took the class within a standard classroom during the district's 5-week summer session.²⁰ The BAU classes primarily relied on traditional teacher-directed instruction, where teachers had latitude in the curriculum and instructional materials for the class.

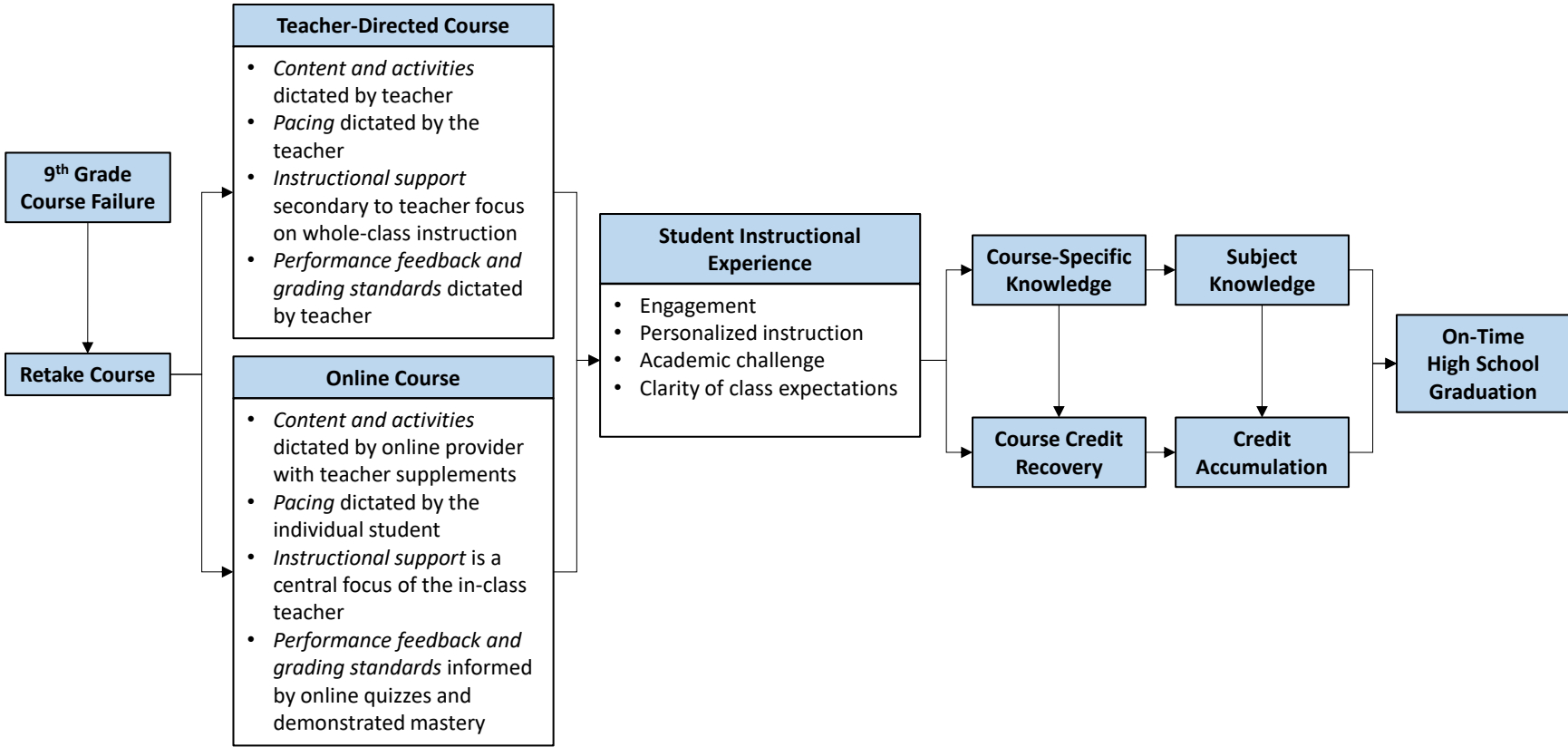
The intervention's theory of action is presented in Figure 1. We hypothesized that the online course would provide students with a different instructional context than the teacher-directed course that could affect the following instructional features:

- Individualized pacing of course content
- Connection of course content to student needs
- Provision of instructional support for student learning
- Provision of more immediate performance feedback to students

¹⁹The study targeted two 9th grade courses with high failure rates in the school district. In the school district, students typically take a year-long first year Algebra course (Algebra 1) in 9th grade and a two semester-long English courses. English 9A is typically taken in the fall and English 9B is typically taken in the spring. We included English 9A and 9B classes in the study, and in this paper report on them together as English 9.

²⁰In addition to the summer sessions, the study included credit recovery classes during the 2018–19 school year. This paper only reports on the analysis of the summer sessions.

Figure 1. Theory of Action



In turn, we hypothesized that exposure to these features affect the following student experiences in the credit recovery classes:

- Engagement in the class
- Personalized instructional support
- Academic challenge in the class
- Clarity of class expectation

Then these experiences can affect course content knowledge and credit recovery, which can ultimately affect more general subject content knowledge, credit accumulation, and high school graduation.



Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit www.air.org.

MAKING RESEARCH RELEVANT

AMERICAN INSTITUTES FOR RESEARCH
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289 | 202.403.5000
www.air.org

LOCATIONS

Domestic: Arlington, VA (HQ) | Sacramento and San Mateo, CA | Chicago, IL | Indianapolis, IN | Waltham, MA | Rockville, MD | Chapel Hill, NC | Austin, TX

International: Ethiopia | Haiti