Journal of Research on Educational Effectiveness
https://doi.org/10.1080/19345747.2023.2198524

A Multisite Randomized Study of an Online Learning Approach to High School Credit Recovery: Effects on Student Experiences and Proximal Outcomes

Jordan Rickles,[a] Margaret Clements,[a] Iliana Brodziak de los Reyes,[a] Mark Lachowicz,[a] Shuqiong Lin,[a] and Jessica Heppen[a]

[a]American Institutes for Research, Arlington, Virginia, USA

To link to this article: https://doi.org/10.1080/19345747.2023.2198524

Abstract

Online credit recovery will likely expand in the coming years as school districts try to address increased course failure rates brought on by the coronavirus pandemic. Some researchers and policymakers, however, raise concerns over how much students learn in online courses, and there is limited evidence about the effectiveness of online credit recovery. This article presents findings from a multisite randomized study, conducted prior to the pandemic, to expand the field's understanding of online credit recovery's effectiveness. Within 24 high schools from a large urban district, the study randomly assigned 1,683 students who failed Algebra 1 or ninth grade English to a summer credit recovery class that either used an online curriculum with in-class teacher support or the school's business-as-usual teacher-directed class. The results suggest that online credit recovery had relatively insignificant effects on student course experiences and content knowledge, but significantly lower credit recovery rates for English. There was limited heterogeneity in effects across students and schools. Non-response on the study-administered student survey and test limits our confidence in the student experience and content knowledge results, but the findings are robust to different approaches to handling the missing data (multiple imputation or listwise deletion). We discuss how the findings add to the evidence base about online credit recovery and the implications for future research.

CONTACT Jordan Rickles jrickles@air.org American Institutes for Research, Arlington, VA, USA.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/19345747.2023.2198524

Article History

Received 18 April 2022

Revised 1 February 2023

Accepted 10 March 2023

Published online 13 April 2023

Keywords

Online learning; credit recovery; high school; randomized controlled trial

Introduction

Even before the coronavirus pandemic pushed online learning to the forefront of education, schools were increasingly turning to online courses as a way to expand course offerings and credit recovery options. Use of online credit recovery grew out of the hope that expanding credit recovery options through online courses will provide students with a more personalized instructional experience and help them get back on track toward graduation (e.g., Atkins et al., 2007; Gemin et al., 2015). With these promises in mind, states and districts invested significant resources into building the infrastructure to offer online credit recovery. As the *Los Angeles Times* editorial board observed in a 2016 editorial, "[Online credit recovery] courses, which have helped boost graduation rates locally and across the country, have grown quickly from a barely known concept a decade ago to one of the biggest and most controversial new trends in education" (*Los Angeles Times* Editorial Board, 2016). Concerns materialized over how much students learn in online courses and potential abuses to raise graduation rates, but the research remains inadequate (Ferdig, 2010; Loewenberg, 2020; Malkus, 2019; Schaeffer & Konetes, 2010; U.S. Department of Education, 2015). The increase in failed courses or missed course offerings during the pandemic will likely amplify the expansion of online credit recovery and the public discourse about its merits, along with the critical need for rigorous evidence about the effective use of online credit recovery for high school students.

Online courses are delivered in varying formats, further complicating how a limited research base can inform practice. Some courses are fully online and completely self-paced; others are hybrid or blended models that combine online learning with face-to-face teacher support for students (Staker & Horn, 2012; Watson & Ryan, 2006). During the 2014–15 school year, 63% of U.S. high schools provided online credit recovery courses and 41% of high schools provided a blended model for credit recovery (U.S. Department of Education, 2018). The promise of online courses for credit recovery lies in features afforded by the technology that, when utilized, can result in courses designed to meet the specific needs of students who are falling behind academically and can provide districts with a more flexible way to expand course availability.

Although many espouse the promise of online courses for credit recovery, and although online credit recovery is used throughout the country (Gemin et al., 2015; Queen & Lewis, 2011), rigorous evidence about its effectiveness is limited (Viano, 2018). Findings from the few studies that focused on the effects of online credit recovery are mixed. One correlational study of online credit recovery in Florida found that students in the online courses were more likely to earn a C or better than students in the face-to-face courses (Hughes et al., 2015). Another study of North Carolina high school students who failed a course found that students who enrolled in online credit recovery were more likely to graduate but had lower test scores than students who repeated the face-to-face version of the course (Viano & Henry, 2020). Based on a study conducted in a large urban midwestern district, Heinrich et al. (2019) concluded that "online course-taking is not benefiting students or reflecting real learning, and some students may even be set back in their learning" (p. 2174). Follow-up research in that same district did find that online course-taking was positively associated with high school graduation (Heinrich & Darling-Aduana, 2021) but negatively associated with longer term labor market earnings (Heinrich & Cheng, 2022).

The only randomized controlled trial of online credit recovery to date found that students who took an online summer credit recovery Algebra 1 course in Chicago were less likely to earn credit and learned less than students who took a face-to-face Algebra 1 course (Heppen et al., 2017). However, exploratory results from the study suggested that students in the online class who received instructional support from an in-class teacher had better academic outcomes than online students who did not receive instructional support (Taylor et al., 2016). In addition, that same study found no statistically significant differences in longer term outcomes, including high school graduation, between students in the online and face-to-face courses (Rickles et al., 2018).

The current study builds on the Chicago study to expand the field's understanding of online credit recovery's effectiveness in three ways: (1) The study focuses on two courses—Algebra 1 and ninth grade English (English 9)—instead of one. (2) The study took place in a different school district and with a different online content provider. (3) In-class teacher support was a more explicit component of the instructional model for the online class. The study took place during the district's 2018 and 2019 summer terms.[1]

This article presents findings on how the online credit recovery classes affected students' in-class experiences and proximal outcomes relative to the business-as-usual (BAU) teacher-directed classes.[2] In particular, we address three main research questions:

1. How did ***students' experiences*** in the online classes compare to those in the teacher-directed classes?
2. How did student ***credit recovery rates*** in the online classes compare to those in the teacher-directed classes?
3. How did students' ***course-specific and general subject content knowledge*** in the online classes compare to those in the teacher-directed classes?

[1] In addition to the summer sessions, the study included a smaller sample of credit recovery classes during the 2018–19 school year. This article only reports on the analysis of the summer terms.
[2] Our analysis was preregistered in the Registry of Efficacy and Effectiveness Studies after the study started but prior to the analysis of outcomes (Registry ID #1917.1v2). The study's research activities were approved by the American Institutes for Research Institutional Review Board (#86436).

In addition to reporting on the primary analysis that focused on the overall average effects of online credit recovery, we summarize results from a series of exploratory analyses that examined treatment effect heterogeneity.

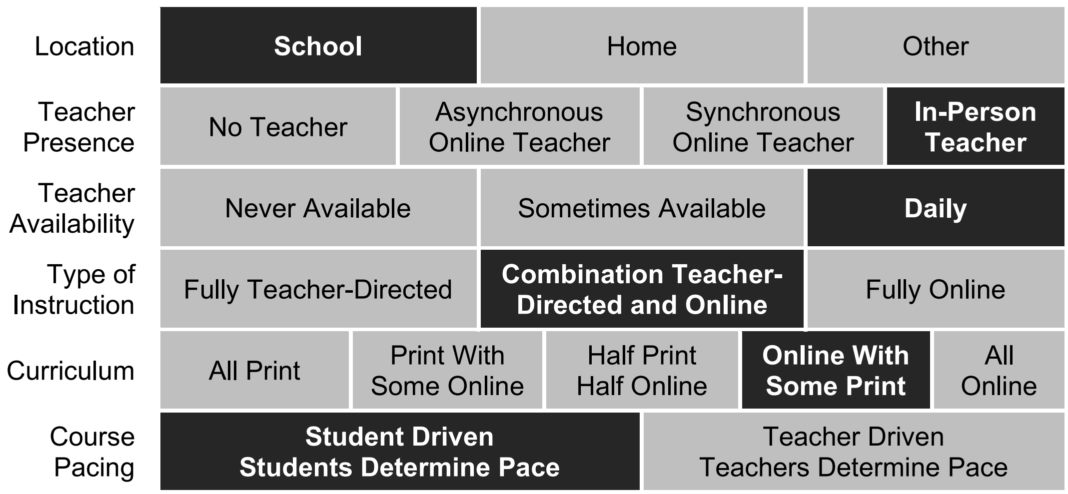
Online Credit Recovery Model and Theory of Action

The intervention for this study was an Algebra 1 or English 9 (first or second semester) online curriculum for the credit recovery course,[3] where an online provider supplied the main course content, and the school provided a subject-appropriate, credentialed in-class teacher who could supplement the digital content with additional instruction. Edgenuity was the online provider for all of the online classes in the study. Figure 1 lays out dimensions on which online classes may vary, highlighting the characteristics of the online learning model we sought to test in this study. The dimensions and figures were adapted from Vanourek (2006) and Watson et al. (2013).

[3] The study targeted two ninth 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 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 report on them together as English 9.

Figure 1

Dimensions of the Online Learning Model



Note. For each dimension, the highlighted aspect represents a key feature of the online learning model for this study.

The rationale for centering a credit recovery course around an online curriculum is based on features afforded by the online program that, if utilized, may result in instructional experiences that meet the needs of academically underserved students. These features include simulations, animations, and interactive tools to promote engagement and support learning; lessons that employ evidence-based models to support learning complex concepts; flexibility that allows students to progress through the course material at their own pace; and providing more immediate performance feedback to students (Bakia et al., 2013; Blackboard K–12, 2009; Dynarski et al., 2008; Kehrer et al., 2013; Kemple et al., 2005; Mayer, 2011; Mayer & Moreno, 2003; Roschelle et al., 2016; U.S. Department of Education, 2009).

The role of the credentialed, in-class teacher provided by each participating school was to monitor students as they worked through the online course and to provide supplemental instructional support targeted to students' needs (Taylor et al., 2016). Teachers could monitor student progress and performance through the online provider's educator dashboard, including information on how much time each student spent in the program and how they performed in each lesson. Although the online curriculum was the core content in these classes, teachers had the flexibility to supplement instruction as they saw fit.

To examine the effectiveness of the online classes, we compared them to the typical teacher-directed credit recovery classes students take at each high school in the study (the BAU condition). Classes in both the online and BAU conditions were expected to be aligned with the district's intended curriculum and cover the same content standards for the course. As with the online classes, a credentialed teacher led the teacher-directed classes. The BAU classes primarily relied on traditional teacher-directed instruction to deliver the course content. In contrast to the online classes, teachers had more flexibility to determine the instructional materials, which were typically paper based. Teachers also had more freedom to determine the pace of student progress through the course content. All classes in both the intervention (online) and BAU (teacher-directed) conditions met for 2.5 hours each day in a standard classroom during the district's 5-week summer session.

The intervention's theory of action is presented in Figure 2. We hypothesized that the online classes would provide students with a different instructional context than the teacher-directed classes and that this different context 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

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

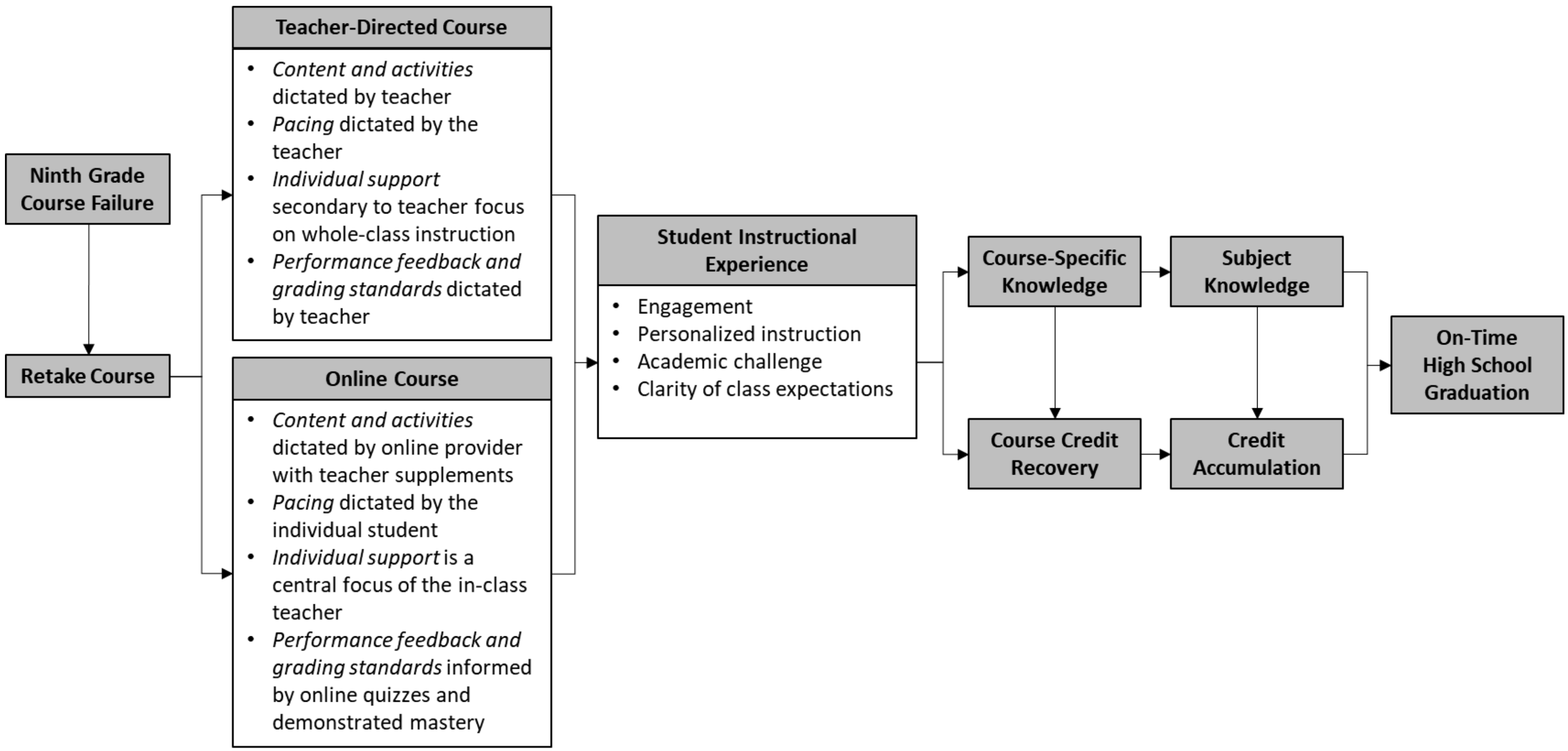
- Engagement in the class

- Personalized instructional support
- Academic challenge in the class
- Clarity of class expectations

These experiences can then affect course credit recovery and subject content knowledge, which can ultimately affect more general subject content knowledge, credit accumulation, and high school graduation. In this article, we examine the course-specific interim outcomes, as well as general subject content knowledge measured a few months after the summer session. As the study continues, we will follow students through 4 years of high school so that we can report on their longer term outcomes in a future article.

Figure 2

Theory of Action



Study Design

In this section, we describe the study sample, then discuss the data collection and measures, and finish with an overview of the analytic approach.

Study Sample

The study took place in Los Angeles high schools that were recruited for the study. To be included in the study, students had to meet the following criteria:

- Entered ninth grade in the 2017–18 or 2018–19 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 or at least one semester of their English 9 course during their first year of high school.
- Enrolled in one of the credit recovery classes that were part of the study at the start of the summer session. For the first summer session (2018), only English 9 classes were in the study. For the second summer session (2019), the study included both Algebra 1 and English 9 classes.
- Not be classified with an English language development (ELD) level of 1, 2, or 3.[4]

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

The analyses in this article 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.[5] In each participating school, half of the classes were online classes and half were teacher-directed classes. 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 semesters of the course the students failed during their ninth grade year. A consort diagram in the Online Supplemental Appendix documents the number of students in the randomized sample and each analytic sample.

Table 1 compares the baseline characteristics of the students in the online and teacher-directed classes. As expected for students requiring credit recovery, students in the study 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) of less than 0.25 standard deviations (a common threshold for baseline equivalence; U.S. Department of Education, 2020). This indicates that the random assignment process successfully resulted in similar students, on average, in each class type.

[5] English 9 includes two courses: English 9A and English 9B. There were 156 students enrolled in more than one class in the same subject (Algebra or English). For these students, we randomly selected which class to include in the analytic sample so that the number of students reflects unique student counts within a subject. There are 54 students in both the Algebra 1 and English 9 analytic samples, so there are 1,683 unique students in the total study sample but 1,737 total observations across the Algebra 1 and English 9 samples.

Table 1

Description of the Student Sample

	Algebra 1			English 9		
Student characteristics	**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 (Level 4 or 5)	18%	19%	−0.03	15%	16%	-0.06
Attendance rate (Grade 9)	92%	92%	0.00	85%	84%	0.04
Average GPA (Grade 9)	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. An omnibus likelihood ratio test of group differences across all covariates was not statistically significant (Algebra 1 $p = 0.663$, English 9 $p = 0.095$).
SMD = standardized mean difference. The SMD was calculated using the Cox index for dichotomous measures and Hedge's *g* for continuous measures. ELD = English language development; GPA = grade point average; SB = smarter balanced scale score standardized on the basis of the districtwide mean and standard deviation; ELA = English language arts.

Data and Measures

We collected the following primary and extant data to address the research questions:[6]

- 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 course content knowledge;
- District extant data, including student background characteristics, eighth grade state test scores, ninth grade academic performance, the final grade students received in the

[6] For primary student-level data collection (student survey and test), students and their guardian were informed that their credit recovery class was participating in a research study and they could opt out of participating in the data collection. Students had to return a form signed by their guardian to opt out of the data collection. The consent form was provided in English and Spanish.

credit recovery course, and student scores on the district-required PSAT fall administration for tenth graders.

Student Survey

The study team administered the student survey in class 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). The surveys included a series of statements about the class and asked students to report how much they agree with each statement, on a 4-point scale, for their credit recovery class.[7] Most of the survey items were adapted from items used in a study of student engagement (Skinner et al., 2009) or the earlier credit recovery study that took place in Chicago (Heppen et al., 2017). With our final sample, including the summer and school-year study classes, 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 actions to learn in class.
- *Emotional engagement* measures the extent to which students felt enthusiasm and enjoyment in class.
- *Personalized instruction* measures the extent to which students thought their teacher provided them with additional instruction when they needed help.
- *Academic challenge* measures the extent to which students thought the class was challenging.

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

- *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 analyzed factor scores standardized based on the total sample mean and standard deviation. More information about each construct is reported in the Online Supplemental Appendix (Table S.1).

Student Test

The student test was administered in a 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 mathematics NAEP tests. We selected items that cover the range of topics typically taught in a first-year algebra course, including some items that cover important prealgebra 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.[8] The test included items about two literary texts and items about 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 (separately for each subject).[9]

[8] For the English test, we used reading passages and items from the Ohio state assessments instead of the NAEP or California's state assessment because Ohio had more publicly available items to choose from and Ohio's state content standards, like California's, were based on the Common Core State Standards.

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

Final Course Grade

The district provided us with the final course grade students received in their credit recovery class. 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 either because they dropped the class or took an incomplete. We coded all students without a final grade as not passing the class during the summer.

PSAT Test Score

The district provided us with PSAT math and reading scores for students who took the PSAT fall administration that the district requires for all tenth graders. Among our student sample, 85% of the Algebra 1 students (84% for online classes and 85% for teacher-directed classes) and 82% of the English 9 students (81% for online classes and 83% for teacher-directed classes) had PSAT scores. While not directly aligned with Algebra 1 and English 9 content, we used the PSAT test scores to examine whether online credit recovery had an effect on more subject-general content knowledge.

District Extant Data

For all students in the study, the district provided us with 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.

Impact Analysis

Our primary analysis targets the individual-level average effect for the finite population of sites in the study sample (Miratrix et al., 2021), and we analyzed the data based on the type of class to which students were randomly assigned so that we could estimate the intent-to-treat (ITT) average effect.[10] We conducted the analysis separately for Algebra 1 and English 9. We used a linear regression model for the survey measure and test score outcomes, and a logistic regression model for whether students passed or did not pass the class. The linear regression model takes on the following form, with a vector of student background characteristics (***X***) as listed in Table 1 and a vector of fixed effects for the randomization blocks (***B***), which account for school (and cohort for English 9):[11]

$$Y_i = \beta_0 + \beta_1 T_i + \boldsymbol{X}_i' \boldsymbol{\beta}_x + \boldsymbol{B}_j' \boldsymbol{\gamma}_j + e_i$$

The parameter of primary interest is β_1, which represents the precision-weighted average treatment effect. The logistic regression model takes the same form but with a log-linear transformation.

Missing data are a potential concern for the Grade 8 state assessment scores (covariates) and the student survey and test (outcomes). In particular, the high overall and differential missing data rates for the student survey and end-of-course test are a potential validity threat for those outcome measures. Based on the What Works Clearinghouse thresholds for acceptable attrition bias (What Works Clearinghouse, 2022), the Algebra 1 sample has high attrition even using the optimistic boundary and the English 9 sample has high attrition under the cautious boundary. Tables comparing the total student sample to the sample of students with scores on the study-

[10] Approximately 90% of the students were enrolled in the class they were assigned to.

[11] 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% to 84%, 85% to 89%, 90% to 94%, and 95% to 100%.

developed end-of-course test and PSAT are provided in the Online Supplemental Appendix (Tables S.2 and S.3). Student characteristics for the sample of students with survey data are nearly identical to the sample of students with test data, given that the survey and test were administered at the same time. Among students with observed outcome data, the treatment and control groups have similar background characteristics (i.e., standardized mean differences below 0.25), which suggests that differential attrition bias may be minimal. However, statistically significant differences in background characteristics existed between students missing outcome data and those with observed outcome data. In particular, students missing survey and end-of-course test scores had significantly lower ninth grade attendance rates, on average, than students with observed outcome data, and students missing PSAT scores had, on average, significantly lower ninth grade attendance rates and lower ninth grade GPA.

To account for potential bias due to missing data, we used multiple imputation with chained equations ($M = 20$). We ran separate imputation models by subject and treatment condition. The imputation models used the predictive mean matching method ($k = 3$) and included all covariates, randomization blocks, and student outcomes.[12] As a robustness check, we estimated treatment effects using a complete case analysis (i.e., listwise deletion). Results from this analysis are presented in Table S.4 in the Online Supplemental Appendix. The direction and statistical significance for all but one effect estimate were the same regardless of whether multiple imputation or listwise deletion was used to handle missing data. The one exception was for the Algebra 1 end-of-course test, where the effect estimate was not statistically

[12] In our preregistered analysis plan, we said we would use multivariate normal regression to multiply impute missing data. That approach produced data with extreme outliers and exacerbated the variance of some measures. As a result, we decided to use predictive mean matching for multiple imputation, which produced more plausible imputed values and distributions.

significant for the main analysis based on multiple imputation but was significant when using listwise deletion.

Exploratory Analysis

We conducted two types of exploratory analyses to examine treatment effect heterogeneity. First, we tested whether average effects varied across sites, defined as school-by-cohort combinations, because the effectiveness of the online class might depend on factors unique to the instructional environment (including the teacher). Second, we tested whether a student's prior academic performance and engagement moderated the effect of online credit recovery because the online class might be more effective for students who enter credit recovery with a relatively strong academic foundation.

For the analysis of between-site heterogeneity, we estimated site-specific treatment effects by adding site-by-treatment interaction terms (fixed effects) to the main impact model specification. We then meta-analyzed the site-specific treatment effect estimates and their standard errors to estimate the between-site treatment effect variance.

For the student-level moderator analysis we focused on three dichotomous student-level indicators: (1) relatively high versus low achievement on the subject-specific eighth grade state assessment, where higher achievement was defined as scoring better than 0.50 standard deviations below the district average; (2) relatively high versus low course performance in the first year of high school, where higher performance was defined as having a GPA above 1.50; (3) relatively high versus low school attendance in the first year of high school, where higher attendance was defined as having an attendance rate above 90%. For the first two indicators, the cut-points split the student sample approximately in half, reflecting the fact that most students in credit recovery enter with low academic performance. For the third indicator, approximately

40% of the student sample is classified in the higher attendance group. We estimated separate moderation models for each dichotomous indicator, where the main impact model was altered to include an interaction term between the treatment indicator and the moderator.

Results

In this section, we briefly describe how students assigned to an online class progressed through the online curriculum to help contextualize the findings and then turn to the three research questions. To present the primary impact analysis results, we first focus on the Algebra 1 classes and then focus on the English 9 classes. We conclude the section with a summary of our exploratory analysis of effect heterogeneity.

In presenting the overall average effect estimates, we provide the 95% confidence interval and 90% confidence interval around the point estimate. The 95% confidence interval aligns with the traditional null hypothesis statistical test, where an interval that intersects with zero implies failure to reject the hypothesis of no effect (with $\alpha = .05$). The 90% confidence interval aligns with the two one-sided tests of equivalence (Barker et al., 2002), where the interval represents the range of differences within which we cannot reject the hypothesis of group equivalence (with $\alpha = .05$). If the threshold for a substantive difference falls outside the 90% confidence interval, we have statistical evidence to conclude that the two groups are substantively equivalent given that threshold. The study was designed to test for significant differences between groups, and therefore we prioritize inferences from that test in our discussion. Presenting the 90% confidence interval provides additional context for how to think about the uncertainty in the effect estimates. Specifically, the 90% confidence interval informs the extent to which we should be cautious to conclude that "no significant difference" implies that the two groups have similar average outcomes.

Online Course Progression

To properly contextualize the effects of online credit recovery, it is helpful to know how much exposure students in the online course had to the online curriculum. For both Algebra 1 and English 9, active time in the online program and online course progression fell below expectations. While there are no formal guidelines for time spent in the online program, the online provider suggested that most students will need approximately 40 hours to get through the content. For students assigned to an online class, only 10% of Algebra 1 students and 12% of English 9 students spent at least 40 hours in the online program content, with the average student spending about 24 hours in the program. Relatedly, only 28% of Algebra 1 and 34% of English 9 students assigned to an online class completed at least three fourths of the online course, with the average student completing 51% of the course content.[13]

Algebra 1 Results

(RQ1) Student Experiences

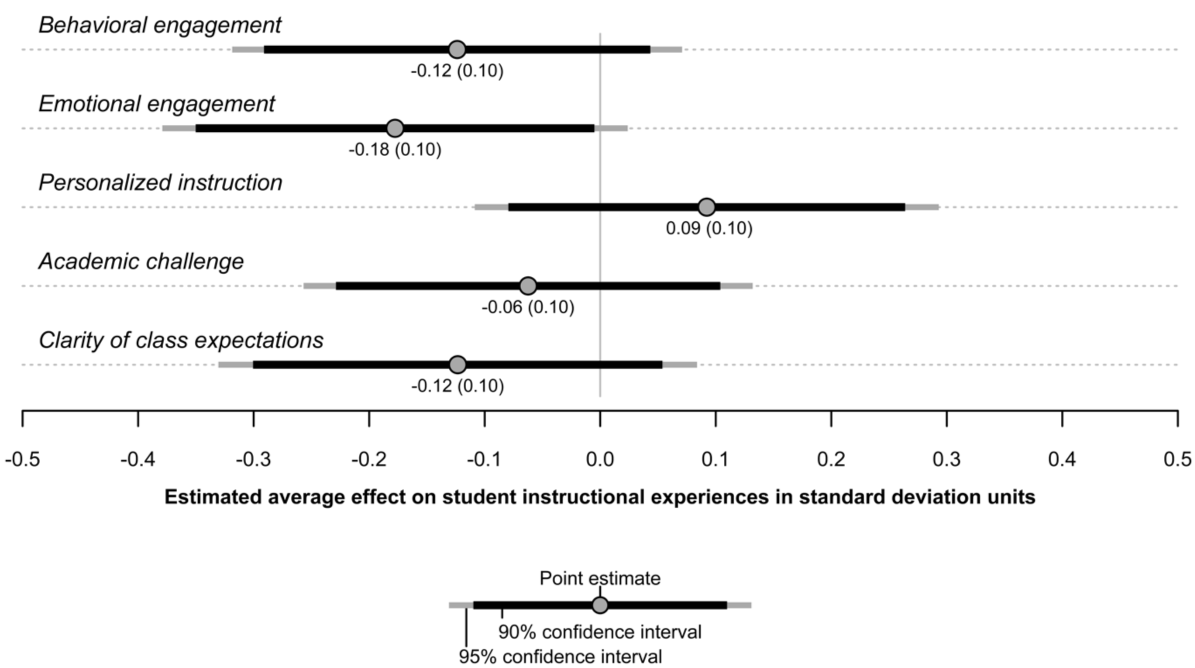
Across the five student experience measures, we found no statistically significant differences between the instructional experiences reported by algebra students assigned to an online class and those of students assigned to a teacher-directed class, on average (see Figure 3). However, the 90% confidence intervals are too wide to conclude that students in the online and teacher-directed classes had substantively equivalent experiences. For example, the lower-bound 90% confidence interval for the effect on behavioral engagement extends to about −0.30 standard deviations, which means we do not have enough confidence in the point estimate to reject a

[13] To examine whether the ITT treatment effect estimates might misrepresent the average effect for a student who completed a significant portion of the online course, we conducted a supplemental analysis to estimate the treatment-on-the-treated (TOT) average effects using a two-stage least squares model with random assignment as the instrument. The TOT estimates based on different thresholds for "treated" are presented in the Online Supplemental Appendix (Table S.5).

hypothesis that the true difference is no larger than 0.30 standard deviations. With the exception of personalized instruction, average differences in student-reported experiences ran counter to the theory of action.

Figure 3

Estimated Differences in Student Instructional Experiences Between Online and Teacher-Directed Algebra 1 Classes



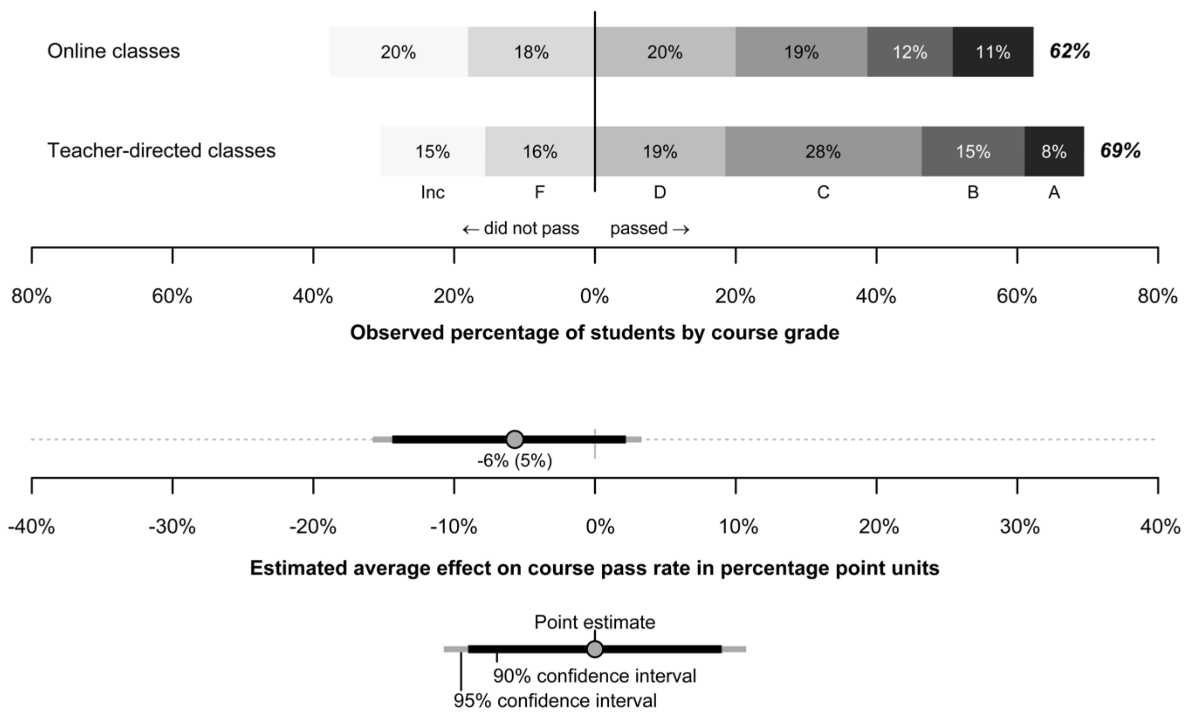
Note. The 95% confidence interval visualizes the null hypothesis test ($\alpha = 0.05$), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence ($\alpha = 0.05$), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. Standard errors are reported in parentheses. $N = 305$ students in the online classes and 308 students in the teacher-directed classes.

(RQ2) Credit Recovery Rates

The average credit recovery rate in the Algebra 1 online classes was 7 percentage points lower than in the teacher-directed classes (62% vs. 69%), with the full grade distribution presented in the top panel of Figure 4. However, the model-estimated treatment effect is −6 percentage points and is not statistically different from zero (see bottom panel of Figure 4). The estimated effect of −6 percentage points represents an effect size of approximately −0.15 based on a Cox index transformation. While the estimated effect is not statistically significant, one should not conclude that the two types of classes had "equivalent" credit recovery rates, because the 90% confidence interval spans beyond a plausible equivalence threshold (e.g., −10 percentage points).

Figure 4

Estimated Differences in Credit Recovery Rates Between Online and Teacher-Directed Algebra 1 Classes



Note. The effect on the course pass rate was estimated with a logistic regression model. We display the effect in percentage points (bottom panel) relative to the observed pass rate in the teacher-directed classes. The 95% confidence interval visualizes the null hypothesis test ($\alpha = 0.05$), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence ($\alpha = 0.05$), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. The standard error is reported in parentheses. $N = 305$ students in the online classes and 308 students in the teacher-directed classes.
Inc = incomplete or no final grade assigned.

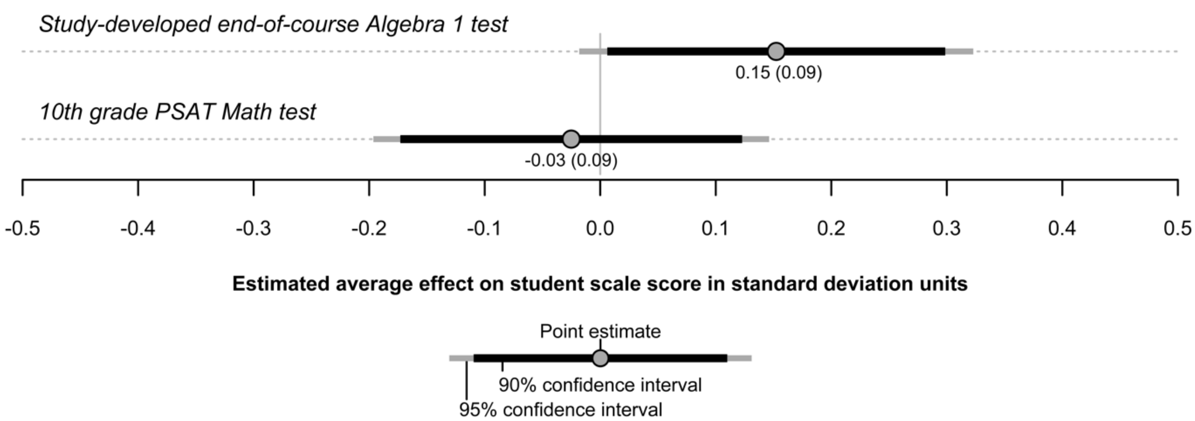
(RQ3) Content Knowledge

The estimated average effect on student content knowledge is presented in Figure 5. On the end-of-course Algebra 1 test, students in the online classes scored 0.15 standard deviations higher, on average, than students in the teacher-directed classes. The estimated difference, however, is not statistically significant.[14] Perhaps more importantly, in both the online and teacher-directed classes, students only answered about a third of the test questions correctly, on average. Any differences in Algebra 1 content knowledge did not translate over to more general math content knowledge, as measured by the PSAT. The average difference in PSAT performance between the two types of classes was only −0.03 standard deviations, which was not statistically significant.

[14] The finding about student performance on the algebra test is sensitive to how we handle missing test score data in the analysis (missing for 34% of the students). For our primary analysis that uses multiple imputation for missing data, the estimated effect is 0.15 ($p = 0.08$), but if we conduct a complete case analysis that drops observations with missing data the estimate effect is 0.21 ($p = 0.03$). For all other outcomes, the statistical significance test inference was the same when using multiple imputation or a complete case analysis. See Table S.4 in the Online Supplemental Appendix for a comparison of the results from the multiple imputation and complete case analyses.

Figure 5

Estimated Differences in Student Content Knowledge Between Online and Teacher-Directed Algebra 1 Classes



Note. The 95% confidence interval visualizes the null hypothesis test (α = 0.05), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence (α = 0.05), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. Standard errors are reported in parentheses. *N* = 305 students in the online classes and 308 students in the teacher-directed classes.

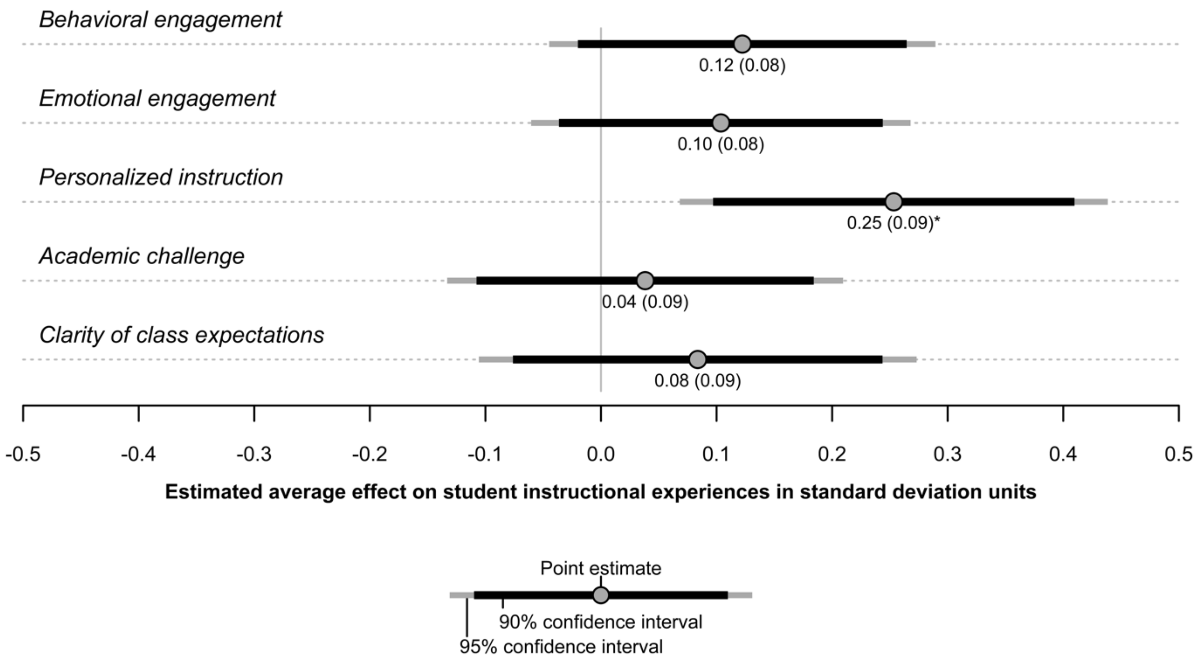
English 9 Results

(RQ1) Student Experiences

The student survey responses suggest that students, on average, had more positive experiences in the English 9 online classes than in the teacher-directed classes (see Figure 6). However, the estimated differences in student experiences were not statistically significant, except for personalized instruction. The average degree of personalized instruction was 0.25 standard deviations more prevalent in the online classes than in the teacher-directed classes.

Figure 6

Estimated Differences in Student Instructional Experiences Between Online and Teacher-Directed English 9 Classes



Note. The 95% confidence interval visualizes the null hypothesis test (α = 0.05), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence (α = 0.05), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. Standard errors are reported in parentheses. $N = 564$ students in the online classes and 560 students in the teacher-directed classes.
*Average difference between the online and teacher-directed classes is statistically significant (p value < 0.05).

(RQ2) Credit Recovery Rates

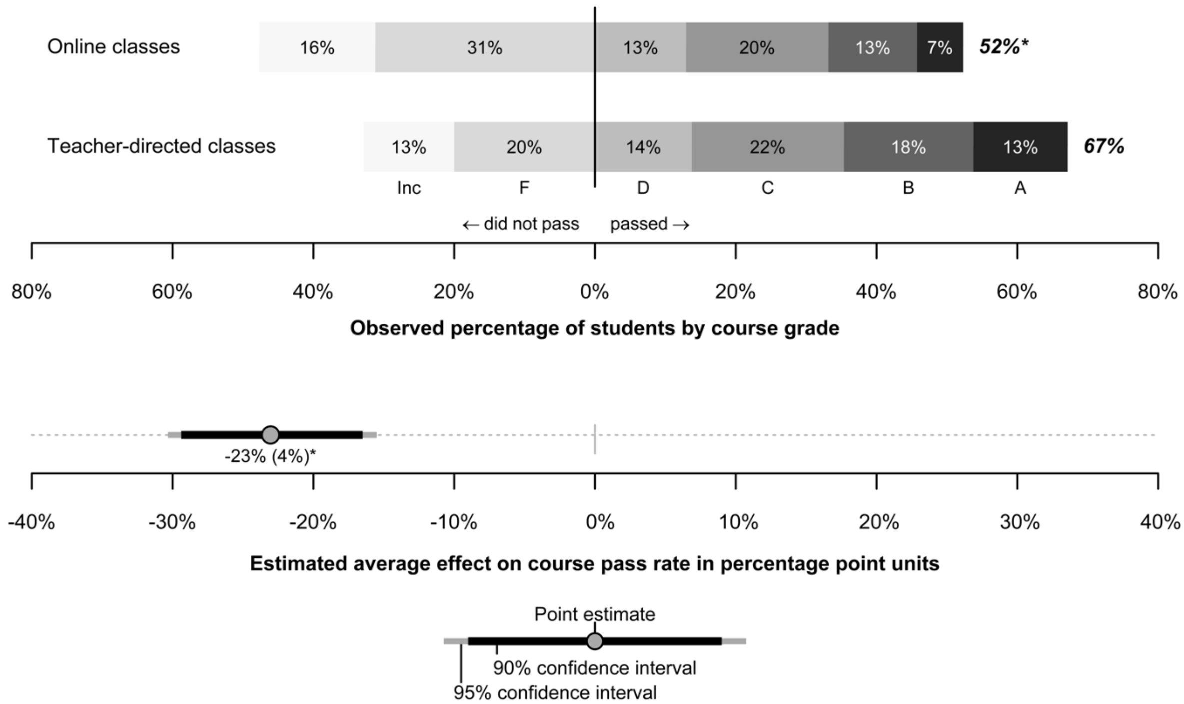
Despite a significant difference in student-reported personalized instruction, credit recovery rates were significantly lower in the English 9 online classes than in the teacher-directed classes (see Figure 7). Based on the observed course grades for the study sample, the credit recovery rate for students assigned to an online class was 15 percentage points lower than for students assigned to a teacher-directed class (52% vs. 67%). The model-estimated average treatment effect indicates that taking the online class had an even worse effect on passing the English 9 course: a statistically significant estimated effect of −23 percentage points. This difference represents an effect size of approximately −0.58 based on a Cox index transformation.

The large negative effect for English 9 was likely driven by unintended shifts in the grading criteria teachers used to determine final course grades. Based on responses to a survey we administered to all teachers in the study, English 9 teachers of the online and teacher-directed classes said they emphasized different criteria when determining final course grades. In particular, tests and quizzes accounted for 67% of the final grade in the average online class and only 40% in the average teacher-directed class. The increased emphasis on tests and quizzes was paired with a decreased emphasis on class assignments, which accounted for 18% of the final grade in the average online class and 38% in the average teacher-directed class. These differences in grading criteria were statistically significant, and there was a small non-significant change in the emphasis on behavior-related criteria. A likely explanation for the shift in grading is that teachers of the online classes utilized the overall grade students received within the online program, which is about 70% test or quiz based and 30% assignment based when determining a student's final grade for the course.[15]

[15] Per district policy, the classroom teacher in both the online and teacher-directed classes have "ownership" of the gradebook and can determine student grades as the teacher determines is best. Teachers of the online class could use the grade produced by the online program as they saw fit. Anecdotally, we observed some confusion among teachers about the latitude they had to determine final grades in the online classes, with many teachers thinking they had to base a student's final course grade on the student's grade in the online program.

Figure 7

Estimated Differences in Credit Recovery Rates Between Online and Teacher-Directed English 9 Classes



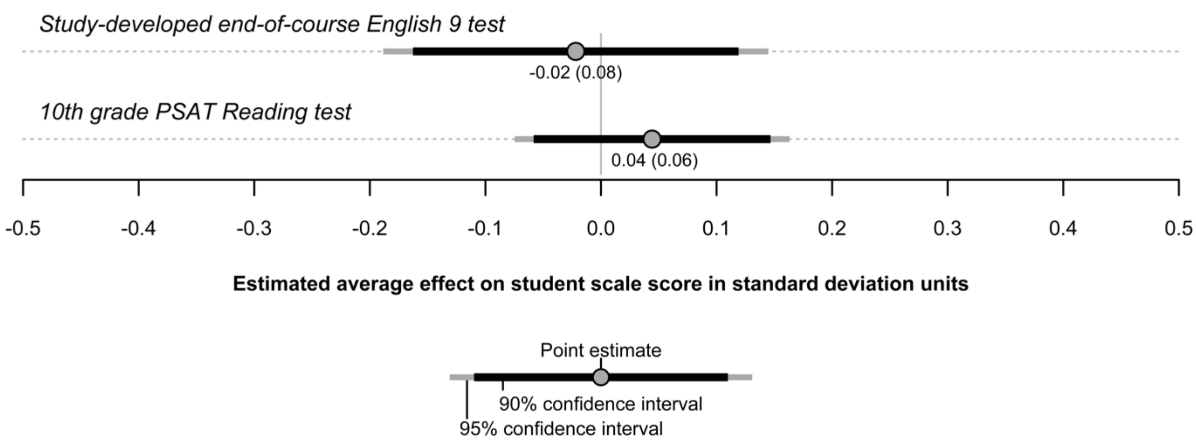
Note. The effect on the course pass rate was estimated with a logistic regression model. We display the effect in percentage points (bottom-right panel) relative to the observed pass rate in the teacher-directed classes. The 95% confidence interval visualizes the null hypothesis test ($\alpha = 0.05$), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence ($\alpha = 0.05$), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. Standard errors are reported in parentheses. $N = 564$ students in the online classes and 560 students in the teacher-directed classes. Inc: incomplete or no final grade assigned.
*Average difference between the online and teacher-directed classes is statistically significant (p value < 0.05).

(RQ3) Content Knowledge

The estimated average effect on student content knowledge is presented in Figure 8. For both tests, average performance was not statistically different in the online and teacher-directed classes. On the end-of-course English 9 test, students in the online classes scored 0.02 standard deviations lower, on average, than students in the teacher-directed classes. On the more general reading PSAT, students in the online classes scored 0.04 standard deviations higher, on average, than students in the teacher-directed classes.

Figure 8

Estimated Differences in Student Content Knowledge Between Online and Teacher-Directed English 9 Classes



Note. The 95% confidence interval visualizes the null hypothesis test (α = 0.05), where intervals that intersect with zero indicate that the effect is not statistically significant. The 90% confidence interval visualizes the two one-sided tests of equivalence (α = 0.05), where intervals represent the range of group differences within which we cannot conclude the groups are statistically equivalent. Standard errors are reported in parentheses. N = 305 students in the online classes and 308 students in the teacher-directed classes.

Exploration of Treatment Effect Heterogeneity

While we found no significant average effects on student content knowledge, the non-significant average effects may mask meaningful effects across instructional situations or for certain types of students. For our exploration of whether the average effect of the online credit recovery classes differed across instructional situations, we tested whether there was significant variation in average effects across school-by-cohort pairs. For both subjects and both content knowledge outcomes, the estimated between-site effect variance was close to zero and not statistically significant. We conducted an analysis of student-level moderators to test whether the online classes had a differential effect based on a student's academic foundation coming into credit recovery. Across the three moderators we tested for each of the two course subjects and two content knowledge outcome measures, results from the moderator analysis do not point to a consistent relationship between a student's academic foundation and the effect of online credit recovery.

A summary of the moderation model results is provided in Figure 9, which displays the model-estimated effect for students with the relatively higher prior performance indicator and the effect for those with the relatively lower prior performance indicator. In all but one of the 12 moderator estimates, differences in the effect among students with higher versus lower prior performance were not statistically significant. For the analysis of the algebra end-of-course test, however, there was a consistent trend across the three prior performance indicators that the online classes had a larger positive effect among students coming in with higher prior performance. For each of the prior performance moderators, the Algebra 1 online classes had a statistically significant positive effect on Algebra content knowledge for students entering the credit recovery course with relatively higher performance, though the differential effect between higher and lower prior performance was only statistically significant for the difference between students with higher versus lower Grade 9 GPA: an estimated average effect of 0.34 standard deviations for higher GPA students versus an average effect of −0.06 standard deviations for lower GPA students. This differential effect remains statistically significant after applying a Benjamini-Hochberg adjustment (Benjamini & Hochberg, 1995) to account for the multiple moderator tests we conducted. The lack of statistically significant results for the other moderator tests could reflect the limited statistical power we had to identify significant moderator effects.[16] Either way, the positive effects for the Algebra 1 end-of-course test among higher prior performance students did not translate to significant positive effects on the broader PSAT, and there were no statistically significant effects for the English 9 online classes.

[16] With our realized sample sizes and outcome variances, the minimum detectable moderator effect (with 80% power) ranged from 0.22 to 0.62 across outcomes and moderator groups.

Figure 9

Estimated Average Effect of the Online Credit Recovery Classes on Student Content Knowledge, by Academic Foundation Moderators

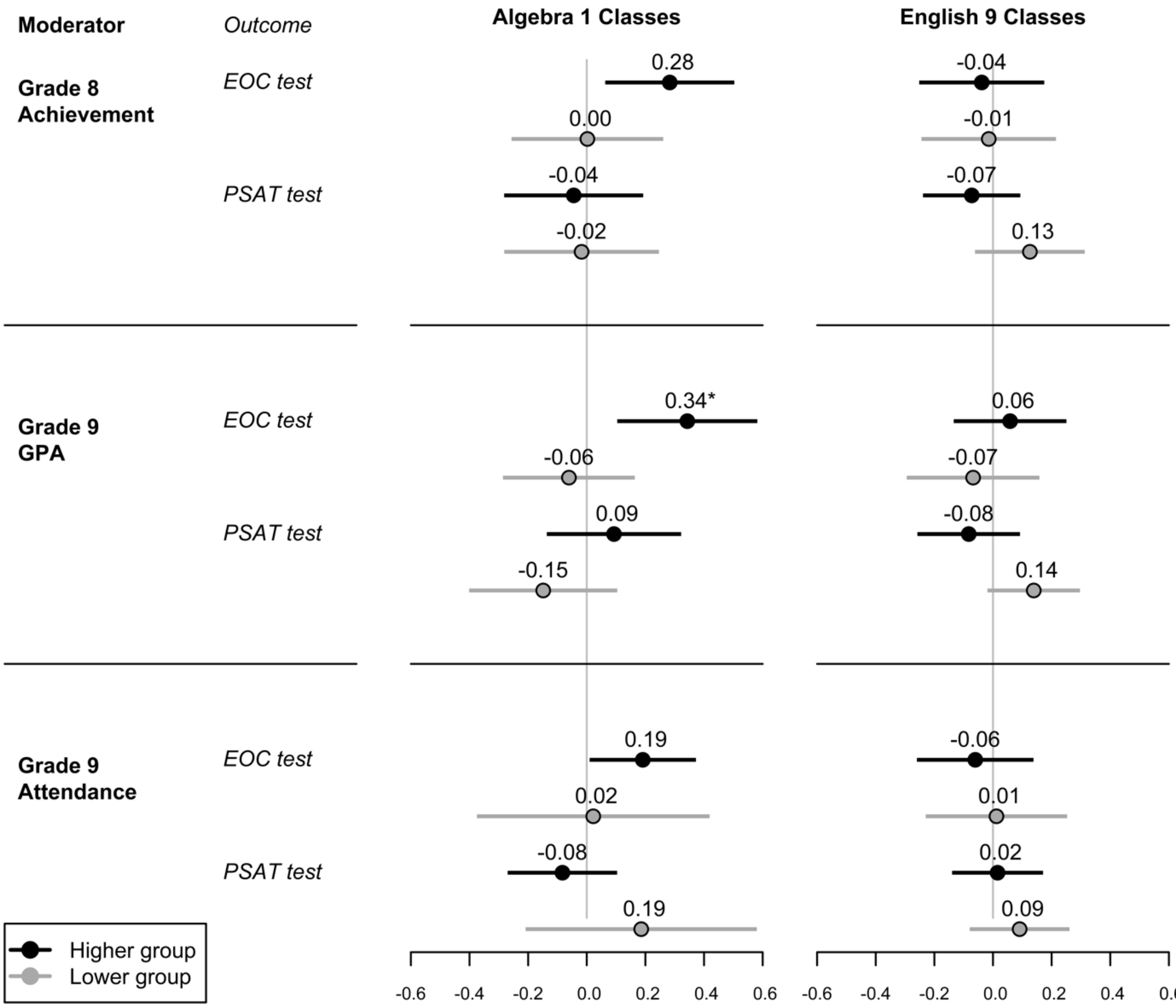


Figure 9. Estimated average effect of the online credit recovery classes on student content knowledge, by academic foundation moderators. *Note.* Effect estimates are in standard deviation units. The horizontal bars represent the 95% confidence interval, where intervals that intersect with zero indicate that the estimated effect is not statistically significant from zero. For Algebra 1, the achievement moderator is based on the Grade 8 mathematics state test score and the test outcomes are the study-developed algebra test and PSAT math. For English 9, the achievement moderator is based on the Grade 8 English language arts state test score and the test outcomes are the study-developed English test and PSAT reading. EOC = end-of-course test.
*Difference between the average effect in the higher versus lower group is statistically significant (p value < 0.05).

Discussion

The Algebra 1 and English 9 results do not provide a consistent story about the relative effectiveness of online credit recovery, but they do inform the evidence base about online credit recovery and have implications for future research. Broadly speaking, the findings do not support the public perception that online courses are easier to pass or that students learn less in these classes. On the contrary, when compared to a school's typical teacher-directed credit recovery course, the Algebra 1 and English 9 findings suggest that students learn about the same, though the analysis lacks enough precision to confidently conclude students demonstrated similar content knowledge in both types of credit recovery classes. For Algebra 1, there is some evidence that students coming into credit recovery with relatively higher prior academic performance may learn more in the online class than in the teacher-directed class. For English 9, the findings suggest that students are less likely to pass an online course. This is not to say that concerns about online credit recovery are necessarily unfounded but to say that the online credit recovery landscape is, like most issues in education, complicated.

Given these complications, it is important to consider this study's findings within the context of the online model implemented. First, we studied an online course that utilized a single online provider, Edgenuity, where student interaction with the online program was suboptimal. Data from the online provider indicate that most students did not spend the recommended time online and completed less than half of the online course. It is not clear how the results would generalize if students spent the recommended time with the online program. Furthermore, it is not clear how this level of engagement with and progression through the online course content compares to how students engage with and progress through the course content in a typical teacher-directed credit recovery class. Future research could seek to disentangle the potential educational value of the instructional model (online vs. teacher directed) from the degree of engagement or take-up. Second, we studied an online model where the primary curriculum was online but there was a credentialed in-class teacher to support student learning, or even augment it with their own instruction. Findings about the instructional features (not presented in this article) and student experiences suggest that we tested a limited version of an online model with in-class instructional support. In particular, we intended teachers in the online classes to provide more instructional support and more frequently supplement the online curriculum to better address their students' needs than occurred during the study. A study based on a well-implemented blended course may find different results on student outcomes. In addition, results from this study may not generalize to a virtual course with little to no instructional support or to a course that uses a different online provider.

It is important to consider the features of the online model we tested not just in relation to interpreting the estimated effects on proximal student outcomes, but also in relation to how much it costs to implement online credit recovery. The potential for cost savings could be one reason school districts consider using online courses, and null effects for online credit recovery may be interpreted as a net-positive if online credit recovery costs less than the teacher-directed option. For example, in their study of online credit recovery, Heinrich and Darling-Aduana (2021) reported that online credit recovery cost the district in their study about half what it cost to provide credit recovery in a traditional classroom. However, such cost savings are most likely to occur when the online courses support larger classes and/or do not require a subject-specific credentialed teacher in the classroom, as was the case in the Heinrich and Darling-Aduana study. In a cost analysis for our study (Atchison et al., 2020), we found that it cost the district about

25% more per student to implement the online classes than the teacher-directed classes.[17] This was primarily because both types of classes in the study required a credentialed in-class teacher and had similar class sizes, but the online classes had the additional cost of the online program licenses. We also found, however, that if you account for the amount of time teachers spent beyond their contracted hours working on the class (e.g., lesson planning, developing materials, grading), the total costs were lower in the online classes than the teacher-directed classes. The total cost savings were particularly pronounced for English 9 classes.

Two other important contextual factors are the setting and student population for this study. We focused on students who retook an Algebra 1 or English 9 course during the summer between their first and second year of high school. These students may be a little more motivated than their peers who failed and did not try to recover the credit during the summer. But they also may face less immediate pressure to pass than the eleventh and twelfth grade students who take online courses to get the credits required for graduation. In talking to school administrators and teachers for our study, for example, some expressed concerns that many of the first-year high school students did not yet have the self-regulatory skills and sense of urgency to succeed in the online course, and that the online course would be better suited for the older high school students. In addition, the study of online education by Heinrich et al. (2019) found that younger high school students (ninth and tenth graders) were less compatible with the online course-taking system because they were less engaged in the online program and the effect on future course

[17] For the cost analysis, we used the ingredients approach, as developed by Levin et al. (2018), to calculate the costs of providing the credit recovery classes. The approach involved identifying the comprehensive list of "ingredients"—personnel and non-personnel resources such as instructor time during and outside of class, computers, and textbooks—associated with providing credit recovery (both online and teacher directed). We collected data from several sources: teacher weekly logs (to get at time spent on class related activities outside of the school day, like planning, grading, and time spent supporting students), end-of-the-course teacher survey (to capture time spent on professional development, and use of materials, books, and equipment) and interviews with school administrators (to assess differences in administrative time required for setting up and managing the two types of classes).

performance was worse for them than for older students. Our study cannot directly speak to this concern, though our data do demonstrate that at least some first-year high school students succeed in online credit recovery. While one should be cautious about extrapolating our findings to distinctly different settings and populations, one should also be cautious to not default to broad assumptions about student capabilities.

The study's limitations spotlight the paucity of research on the effects of online credit recovery. This study and the study in Chicago (Heppen et al., 2017) represent the only two experimental tests of online credit recovery. Both studies focused on students taking credit recovery during the summer between their first and second year of high school but used different online providers and tested different online models. In Chicago, there was an online instructor and an in-class "mentor" that did not have an explicit expectation to provide instructional support, while in Los Angeles there was no online instructor and the in-class teacher was expected to provide instructional support. It is impossible to say, however, whether the different findings regarding Algebra 1 test scores (a statistically significant negative 0.19 effect size in Chicago vs. a positive 0.15 but not statistically significant effect size in Los Angeles) are due to differences in the provider or instructional model. What we can say from these divergent results is that there is more to learn, and decision-makers should be cautious about expanding online credit recovery without studying its effects within their local context and implementation model.

Open Research Statements

Study and Analysis Plan Registration

This study is registered on the Registry of Efficacy and Effectiveness Studies (REES Registry ID 1917).

Data, Code, and Materials Transparency

The data, code, and materials that support the findings of this study are not publicly available.

Design and Analysis Reporting Guidelines

This submission included a completed copy of the JREE Randomized Trial Checklist.

Transparency Declaration

The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Replication Statement

This manuscript reports an original study.

Acknowledgments

We 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 study, particularly Lindsay Marczak; and the school leaders, teachers, and students who participated in the study. The opinions expressed in this paper are those of the authors and do not represent the views of the Los Angeles

Unified School District or Edgenuity. We would also like to thank Ann-Marie Faria and Whitney Cade for their feedback on earlier drafts of the article, as well as the anonymous reviewers for the journal.

Open Scholarship

This article has earned the Center for Open Science badges for Preregistered through Open Practices Disclosure. The materials are openly accessible at https://sreereg.icpsr.umich.edu/sreereg/subEntry/2546/pdf?section=all&action=download

Funding

This study was funded with research grant R305A170152 from the U.S. Department of Education's Institute of Education Sciences. The opinions expressed are those of the authors and do not represent the views of the Institute of Education Sciences or the U.S. Department of Education.

References

Atchison, D., Rickles, J., Clements, P., Brodziak de los Reyes, I., & Heppen, J. (2020). *Online credit recovery: Resources and costs*. American Institutes for Research. https://www.air.org/sites/default/files/Online-Credit-Recovery-Study-Brief-4-Resources-Oct-2020.pdf

Atkins, D., Brown, J. S., & Hammond, A. (2007). *A review of the open educational resources (OER) movement: Achievement, challenges and new opportunities*. Report for the William and Flora Hewlett Foundation.

Bakia, M., Mislevy, J., Heying, E., Patton, C., Singleton, C., & Krumm, A. (2013). *Supporting K–12 students in online learning: A review of online Algebra I courses*. SRI International.

Barker, L. E., Luman, E. T., McCauley, M. M., & Chu, S. Y. (2002). Assessing equivalence: an alternative to the use of difference tests for measuring disparities in vaccination coverage. *American Journal of Epidemiology*, *156*(11), 1056–1061. https://doi.org/10.1093/aje/kwf149

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Blackboard K–12. (2009). *Credit recovery: Exploring answers to a national priority*. Blackboard.

Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., & Smink, J. (2008). *Dropout prevention: A practice guide* (NCEE 2008–4025). National Center for Education

Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Ferdig, R. E. (2010). *Understanding the role and applicability of K–12 online learning to support student dropout recovery efforts*. Michigan Virtual University. http://www.mivu.org/Portals/0/RPT_RetentionFinal.pdf

Gemin, B., Pape, L., Vashaw, L., & Watson, J. (2015). *Keeping pace with K–12 digital learning: An annual review of policy and practice*. Evergreen Education Group.

Heinrich, C. J., & Cheng, H. (2022). Does online credit recovery in high school support or stymie later labor market success? *Journal of Policy Analysis and Management*, *41*(4), 984–1011. https://doi.org/10.1002/pam.22376

Heinrich, C. J., & Darling-Aduana, J. (2021). Does online course-taking increase high school completion and open pathways to postsecondary education opportunities? *Educational Evaluation and Policy Analysis*, *43*(3), 367–390. https://doi.org/10.3102/0162373721993485

Heinrich, C. J., Darling-Aduana, J., Good, A., & Cheng, H. (2019). A look inside online educational settings in high school: Promise and pitfalls for improving educational opportunities and outcomes. *American Educational Research Journal*, *56*(6), 2147–2188. https://doi.org/10.3102/0002831219838776

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. http://doi.org/10.1080/19345747.2016.1168500

Hughes, J., Zhou, C., & Petscher, Y. (2015). *Comparing success rates for general and credit recovery courses online and face to face: Results for Florida high school courses* (REL 2015–095). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://files.eric.ed.gov/fulltext/ED559978.pdf

Kehrer, P., Kelly, K., & Heffernan, N. (2013). Does immediate feedback while doing homework improve learning? In *Proceedings of the Twenty-Sixth International FLAIRS Conference* (pp. 542–545).

Kemple, J., Herlihy, C., & Smith, T. (2005). *Making progress toward graduation: Evidence from the talent development high school model*. Manpower Demonstration Research Corporation.

Levin, H. M., McEwan, P. J., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis.* SAGE Publications.

Loewenberg, D. (2020). A digital path to a diploma: Online credit recovery classes are a lifeline—and ripe for abuse. *Education Next*, 20(1), 50–56. https://www.educationnext.org/files/ednext_XX_1_loewenberg.pdf

Los Angeles Times Editorial Board (2016, June 18). What's really in LAUSD's online credit recovery courses? *Los Angeles Times.* http://www.latimes.com/opinion/editorials/la-ed-credit-recovery-courses-20160619-snap-story.html

Malkus, N. (2019). *Practice outpacing policy? Credit recovery in American school districts*. American Enterprise Institute. http://www.aei.org/wp-content/uploads/2019/11/Practice-Outpacing-Policy.pdf

Mayer, R. E. (2011). *Applying the science of learning.* Pearson.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43–52. https://doi.org/10.1207/S15326985EP3801_6

Miratrix, L. M., Weiss, M. J., & Henderson, B. (2021). An applied researcher's guide to estimating effects from multisite individually randomized trials: Estimands, estimators, and estimates. *Journal of Research on Educational Effectiveness*, *14*(1), 270–308. https://doi.org/10.1080/19345747.2020.1831115

Queen, B., & Lewis, L. (2011). *Distance education courses for public elementary and secondary school students: 2009–10* (NCES 2012–008). U.S. Department of Education, National Center for Education Statistics.

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. https://doi.org/10.3102/0013189X18788054

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, *2*(4). https://doi.org/10.1177/2332858416673968

Schaeffer, C. E., & Konetes, G. D. (2010). Impact of learner engagement on attrition rates and student success in online learning. *International Journal of Instructional Technology & Distance Learning*, *7*(5), 3–9.

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

Staker, H., & Horn, M. B. (2012). *Classifying K–12 blended learning*. Innosight Institute. http://files.eric.ed.gov/fulltext/ED535180.pdf

Taylor, S., Clements, P., Heppen, J., Rickles, J., Sorensen, N., Walters, K., Allensworth, E., & Michelman, V. (2016). *Getting back on track: The role of in-person instructional support for students taking online credit recovery*. American Institutes for Research. https://www.air.org/sites/default/files/downloads/report/Online-vs-F2F-Credit-Recovery.pdf

U.S. Department of Education, Institute for Education Sciences, What Works Clearinghouse (2015). *Dropout prevention intervention report: Credit recovery programs*. Author. https://ies.ed.gov/ncee/wwc/docs/interventionreports/wwc_credit_050515.pdf

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

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.* Author. https://files.eric.ed.gov/fulltext/ED505824.pdf

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2018). *Issue brief: Credit recovery.* Washington, DC: Author. http://files.eric.ed.gov/fulltext/ED591002.pdf

Vanourek, G. (2006). A primer on virtual charter schools: Mapping the electronic frontier. *National Association of Charter School Authorizers Issue Brief*, 10.

Viano, S. L. (2018). At-risk high school students recovering course credits online: What we know and need to know. *American Journal of Distance Education*, *32*(1), 16–26. https://doi.org/10.1080/08923647.2018.1412554

Viano, S., & Henry, G. T. (2020). An evaluation of credit recovery as an intervention for high school students who fail courses (EdWorkingPaper: 20-302). Annenberg Institute at Brown University. https://doi.org/10.26300/kcx1-vk12

Watson, J., Murin, A., Vashaw, L., Gemin, B., & Rapp, C. (2013). *Keeping pace with K-12 online & blended learning: An annual review of policy and practice.* 10 Year Anniversary Issue. Evergreen Education Group.

Watson, J., & Ryan, J. (2006). *Keeping pace with K–12 online learning: A review of state-level policy and practice*. Evergreen, CO: Evergreen Consulting Associates. http://www.kpk12.com/wp-content/uploads/KeepingPace_2006.pdf

What Works Clearinghouse (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE).