Do Early Warning Systems Help High School Students Stay on Track for College? Mixed Methods Evaluation of the Ninth Grade Success Initiative

David S. Knight
University of Washington

Julia Duncheon
University of Washington

Kefi Anderson
Office of the Superintendent of Public Instruction

Matthew Frizzell
Office of the Superintendent of Public Instruction

As the transition point between middle school and high school, ninth grade can either set a student up for long-term success or diminish a student's likelihood of graduating high school altogether. Interventions that can help educators better meet the needs of students during this critical juncture represent powerful levers for driving school improvement. The Ninth Grade Success Initiative is a dropout prevention program, piloted in five Washington State high schools in 2019-20. We use multiple methods to evaluate effects on student outcomes and implementation processes. We find that the program led to improvements in course grades and, to a lesser degree, behavioral outcomes, with little change in student attendance. Data coaches perceived that this program led to more effective targeting of services to higher-need students and better preparation for the COVID-19 transition to virtual learning.

Do Early Warning Systems Help High School Students Stay on Track for College? Mixed
Methods Evaluation of the Ninth Grade Success Initiative *

David S. Knight (corresponding author)
University of Washington
dsknight@uw.edu

Julia Duncheon
University of Washington
duncheon@uw.edu

Kefi Anderson
Office of the Superintendent of Public Instruction
kefi.andersen@k12.wa.us

Matthew Frizzell
Office of the Superintendent of Public Instruction
matthew.frizzell@k12.wa.us

January 2022

Abstract

As the transition point between middle school and high school, ninth grade can either set a student up for long-term success or diminish a student's likelihood of graduating high school altogether. Interventions that can help educators better meet the needs of students during this critical juncture represent powerful levers for driving school improvement. The Ninth Grade Success Initiative is a dropout prevention program, piloted in five Washington State high schools in 2019-20. We use multiple methods to evaluate effects on student outcomes and implementation processes. We find that the program led to improvements in course grades and, to a lesser degree, behavioral outcomes, with little change in student attendance. Data coaches perceived that this program led to more effective targeting of services to higher-need students and better preparation for the COVID-19 transition to virtual learning.

Do Early Warning Systems Help High School Students Stay on Track for College? Mixed

Methods Evaluation of the Ninth Grade Success Initiative

Ninth grade is a critical time in a student's educational trajectory (Allensworth & Easton,

2007). As the transition point between middle school and high school, the freshman year can

either set a student up for long-term success or diminish a student's likelihood of graduating high

school altogether (Neild, 2009). Numerous studies have noted the link between ninth grade

success and high school graduation (McCallumore & Sparapani, 2010; Royster et al., 2015),

noting that students who struggle in that first year are much more likely to leave high school

without a diploma (Cohen & Smerdon, 2009; McSpadden et al., 2008; Warren et al., 2014). In

contrast, as Melissa Roderick and colleagues (2008) noted in their landmark report, students who

stay on track to graduate during their first year in high school—passing all of their freshman year

courses—are significantly more likely to finish high school and matriculate in college.

Subsequent studies have bolstered these finding, suggesting the importance of prioritizing ninth

grade student success to facilitate greater educational opportunity and equity, especially for

students from marginalized populations (Hoffman & Lowizki, 2005; Knight-Manuel et al.,

2019).

This body of work has also identified factors that help ninth grade students succeed, such

as having access to academic support and building meaningful relationships with adults in the

school building (Butts et al., 2005; Roybal et al., 2014). Yet, school-level policies and practices

(e.g., stringent attendance requirements) often undermine students' academic progress,

encouraging student-leaving by making academic success hard to achieve (Brown et al., 2019).

School staff often lack systems and processes to determine which ninth graders are doing well

and which are not, which makes it more likely that students may leave before an adult intervenes

(Rumberger, 2011). Fortunately, new data systems have streamlined the ability of schools to track student progress through data dashboards and progress indicators (Roderick et al., 2011). So called, "early warning" programs leverage student outcome data to help school staff identify individual students who require support. These programs are growing in popularity (Lee & Chung, 2019) and research suggests they show promise (Mac Iver et al., 2017). However, the existing literature is limited in a few ways that the present study aims to address. Few studies have evaluated the impact of early warning systems on student outcomes using a causal framework (Balfanz & Byrnes, 2019). A recent review by Wentworth and Nagaoka (2020) also found that as more schools and districts adopt early warning systems, research is needed to better understand how indicators are used (Davis et al., 2019) and what conditions support their use (Davis et al., 2013).

To these ends, we employ quantitative and qualitative methods in tandem to evaluate the Ninth Grade Success Initiative, a dropout early warning and intervention data system piloted in five Washington state school districts during the 2019-20 school year. The intervention starts with the formation of a Success Team at each participating high school. The Success Team typically includes teachers and other staff who meet multiple times per month to review a student data dashboard. Participating schools are paired with an external coach who helps create the data dashboard and facilitates Success Team meetings. These schools serve populations that are predominantly low-income and of Color and have struggled to increase retention rates for their most vulnerable students. Our quantitative analysis uses a novel approach to addressing the influences of COVID-19 school closure, tracking student outcomes up to March 1 of each year including the intervention year, 2019-20. This approach allows us to examine changes in outcomes in the treatment year, prior to school closures and transitions to virtual instruction. We

also conduct a qualitative analysis of the pilot's implementation process using interview data collected from school coaches. In preview, we find significant impacts on student behavioral outcomes and course passing rates, especially for students classified as English learner or as low-income, but no impact on student attendance outcomes. Results from qualitative analyses highlight several benefits of the pilot for participating schools, as well as school-level conditions that facilitated successful implementation. In the sections below, we first describe the theory of change for the pilot initiative. In subsequent sections, we describe the evaluation methodology, results, and overall conclusions and recommendations for policy.

Intervention and Theory of Change

The Ninth Grade Success pilot initiative supports efforts to improve attendance, achievement, and behavioral outcomes. The goal of the initiative is to provide additional support to grade eight students identified to be at risk of dropping out of high school. Educators use a data dashboard system to identify students who are not making steady academic progress or who are falling behind their peers on these three key outcomes – attendance, behavior, and course performance – as they transition into ninth grade. The dashboard, which is accessible to teachers and coaches in pilot schools, provides information on eighth and ninth grade student attendance, course grades, and behavior referrals.

As part of the pilot initiative in Washington, three instructional coaches work with teachers, the principal, a team lead, and the Ninth Grade Success Team. The Team Lead is a teacher leader who teaches ninth grade and has agreed to serve in a leadership role for the Ninth Grade Success pilot. The Success Team includes all ninth-grade teachers willing to participate in the initiative. Success Team members meet between one and three times per month to review

student data. Instructional coaches are funded through an external organization, Stand for Children, and work directly with teachers.

A summary of the theory of change for the Ninth Grade Success Initiative is shown in Figure 1. The two key drivers of the program are timely supports for attendance, behavior, and course grade data (ABC data) and meaningful team reflection. When these two conditions are in place, a set of four "change ideas" ultimately lead to improved ABC outcomes. These include a Ninth Grade Success Team that meets regularly, access to data visualization, job-embedded coaching, and quarterly meetings called Collaboratives, which provide teachers with opportunities to further explore student data and network with local peers.

Setting, Data, and Analytic Approach

Setting

The study involves five high schools in five different districts located in central Washington state. The pilot high schools include Ellensburg, Grandview, Toppenish, Wahluke, and Moses Lake. Pilot schools serve a higher percent of students who identify as or are categorized as low-income and Latinx compared to the typical Washington state high school. As shown in Table 1, four of the five high schools serve a population where over half of the students are low-income and three high schools, Grandview, Toppenish, and Wahluke, serve greater than 85 percent of students who identify as Latinx. In contrast to the other pilot school districts, Ellensburg enrolls a higher percent of students who identify as white and serves a slightly lower percent of students who are from low-income families and students who are categorized as English learners.

Data

To evaluate the success of the pilot program during the 2019-20 school year, its first year of implementation, evaluators conducted interviews with the three instructional coaches in December 2020 and examined quantitative data from several sources, including the following:

School level data. We requested restricted-use school level data from the Office of the Superintendent of Public Instruction (OSPI) covering student demographics, achievement, attendance, and behavior. Data are disaggregated by grade and by student classification, including household income level and English learner status. These data provide information on school characteristics including grade levels served, total enrollment, and geographic location. Student demographic data include race/ethnicity, gender, household income thresholds, enrollment in special education services, and categorization as an English learner (see Table 1).

As mentioned earlier, we used OSPI data to track three outcomes aligned to the objectives of the pilot, commonly called the ABC outcomes: attendance, behavior, and ninth grade course passing rates. We used student outcome data up to March 1 of each school year so that outcomes for the treatment year, 2019-20 are assessed prior to COVID-19 school closures. We also use outcome data that capture the full school year outcomes. While the full school year data are not available during the treatment year, 2019-20, due to school closures, we use those data to assess the extent to which they align with the March 1 data for each school and school year and to examine any difference between treatment and comparison group schools in the years leading up to the intervention. We defined the ABC outcomes as:

- Attendance: The percent of ninth graders in the participating high school who attended at least 90% of the school day. Students enrolled in the participating high school for fewer than 90 calendar days during the school year are excluded from the analysis, a practice OSPI uses in public data releases (OSPI, 2020).

- Behavior: The percent of ninth graders excluded from the regular school day for disciplinary reasons for three or fewer days during the school year, of those excluded for any duration of time. This measure captures the severity of exclusionary actions for those

students who receive any exclusionary action. To gain further insights, we also examine different ranges of the number of days excluded and the overall percent of students excluded for any amount of time.

- Course passing rates: The percent of first-time ninth graders in the participating high school who earn credit in all attempted courses during their ninth grade year. To earn credit, students must have earned a grade of C or higher. Withdrawals and incompletes are not included in this measure.

As noted, we collected outcomes up to March 1 of each school year so that outcomes for the treatment year, 2019-20 are assessed prior to COVID-19 school closures. Outcome data for the full school year are available for all five pre-intervention years, but not the intervention year. For March 1 outcome data, attendance and course pass rates are available for all five pre-intervention years and the first intervention year, 2019-20. For behavior/discipline data, our outcomes up to March 1 are only available for 2017-18 forward, covering two pre-intervention years and the intervention year.

While we need to use the March 1 outcome data to assess program impacts (since full school year data are not available during the treatment year), we can use the full school year data to determine the extent to which the March 1 outcomes reflect full school year outcomes. For attendance and course pass rates, there are notable (and expected) differences between the full academic year data and the data up to March 1. Both measures are higher for the March 1 outcomes, but we do not find significant differences in those discrepancies across individual schools or between treatment and comparison schools. We also find notable differences for the overall percent of students excluded for any amount of time (an alternate measure of behavior/discipline). However, our preferred discipline measure, the proportion of ninth graders excluded from the school day for disciplinary reasons for three or fewer days during the school year, *of those receiving some sort of exclusionary discipline*, (a measure of discipline severity), is

relatively consistent for data up to March 1 of each school year and for data covering the whole school year. As such, we are able to track all three outcomes from 2014-15 to 2019-20.

Freshman Success Inventory survey. A second source of data came from a survey of Success Team members. The Freshman Success Inventory survey is a progress monitoring data collection tool from the Freshman Success Framework. This survey is administered to Success Team participants at the end of the pilot's first year. The survey asks about the extent to which school principals, Team Leads, and Success Team members from each school district have engaged in particular activities centered around four main areas: setting initial conditions, implementation, communication, and instruction. The survey instrument is adapted from the Network for College Success Freshman On-Track Toolkit, from the University of Chicago School of Social Service Administration. A copy of the survey is included in Appendix Table A1.

Ninth Grade Success Pilot Reflections surveys. Finally, we gathered additional survey data from school staff, referred to as the Reflections survey**.** The Reflections survey is administered to participating school staff at the conclusion of the 2019-20 school year. The survey asks team members and administrators about their experiences during the 2019-2020 school year, including coaching frequency, the role of COVID-19, use of data, barriers to improvement, benefits of the pilot, coaching activities, team collaboratives, data dashboards, and plans to continue using the pilot program in future school years.[1]

Analytic Approach

Quantitative methods. We use survey data from the Freshman Success Inventory survey to assess changes in the conditions for learning in pilot high schools, as perceived by principals,

---

[1] Available here: https://reporting.alchemer.com/r/68766_5f99dbdee465e1.96859149

team leads, and success team members. For each school, we calculate the average survey

responses from all three groups of respondents. Based on the design of the survey, we group

survey items into four context areas, setting initial conditions, implementation, communication,

and instruction. We then compare fall and spring scores for each of the four context areas.

To assess the causal impact of the pilot on the three primary outcomes of interest,

attendance, behavior, and course pass rates described earlier, we compare changes over time in

pilot schools to changes in outcomes in a set of comparison group schools. Changes in outcomes

in pilot schools during the first year of implementation may be related to the intervention, but

other factors may influence estimated effects. For example, Washington state passed several

major school finance reforms in recent years that increased resources available to school

districts. Meanwhile, the COVID-19 pandemic and school closures, which took place halfway

through the first year of implementation, placed substantial strain on students and teachers.

These external factors would bias estimates of program effects based on simple before after

comparisons.

To address concerns about external factors biasing the estimated effects of the pilot, we

use a comparison group of schools that had similar experiences but did not receive the

intervention. These comparison schools serve as a counterfactual for pilot schools. The

counterfactual provides an estimate of pilot school outcomes had they not participated in the

pilot and received the intervention. For the comparison group schools to provide a valid

counterfactual, two conditions must be met (Angrist & Peschke, 2010). First, the two groups of

schools must have parallel trends in outcomes leading up to the intervention. If pilot school

outcomes are decreasing in the years leading up to the first year of intervention, and comparison

group schools are increasing, then the comparison group would not be a valid counterfactual.

Second, we need to rule out any other major events taking place in 2019-20 that may influence

outcomes in the pilot schools, but not in the comparison group schools. We assume that any

major event that took place in 2019-20 (e.g., school finance reforms or COVID-19) affected both

pilot and comparison group schools and so under parallel pre-treatment trends in outcomes, the

comparison schools serve as a relatively valid counterfactual.

To establish a comparison group of schools, we identify all 48 traditional public high

schools located in the same county as one of the pilot schools. Table 2 shows summary statistics

for all Washington high schools, pilot high schools, non-pilot high schools located in the same

three counties as pilot schools, Grant, Kittitas, and Yakima, and all other public high schools in

Washington state. As noted earlier, pilot schools on average serve a higher percent of low-

income students (71%) and students who identify as Latinx (68%), compared to the state

average. Pilot high schools enroll 1,170 students on average, but this figure is skewed upward by

Moses Lake, which enrolls 2,420 students, while Grandview, Ellensburg, and Toppenish each

enroll less than half that amount, and Wahluke High School enrolls less than one-third that

amount. The 48 non-Pilot schools located in the same three counties as pilot schools enroll

significantly fewer students (an average of 403) but have otherwise similar student

demographics.

Among high schools in the same three counties as the pilot schools, only one student

characteristic variable is statistically significant between pilot and comparison group high

school; the percent of students who identify as two or more races is about two percentage points

greater in comparison high schools. As shown in the last two columns of Table 2, all other high

schools have statistically significant differences across a range of student characteristics,

suggesting that these schools may not represent as suitable a comparison group as those high

schools that neighbor the pilot high schools.

To estimate the impact of the pilot, we compare changes over time in outcomes in pilot

schools to that of comparison schools. This analytic approach is referred to by various names,

including a non-parametric difference-in-differences estimator. We estimate the following

model:

$$Y_{it} = \alpha_1 Pilot_i + \sum_{\tau=2015}^{2020} \delta_t * (I_{t=\tau} + Pilot_i{'}\beta_\tau) + X_{it}{'}\gamma + \sigma_c + \varepsilon_{it} \qquad (1),$$

where Y_{it} represents the relevant outcome for high school I in year t, $Pilot_i$ is an indicator for

whether the high school is one of five pilot high schools, δ_t are year indicator variables, X_{it} is a

vector of controls that includes log enrollment, the percent of students classified as low-income

or English learner, and the percent of students receiving special education services, σ_c represents

county fixed effects, and ε_{it} is an error term assumed independently identically distributed. The

causal impact of the pilot program on each of the student outcomes is given by $\beta_{\tau=2020}$, which

estimates the difference in outcomes between treatment and the comparison group schools in the

2019-20 school year, relative to the prior year difference in outcomes between the two groups of

schools (using the 2018-19 school year as the base year for comparison). We use equation 1 to

generate all time-series figures, but for the purpose of reporting regression coefficients (in

Appendix Table A1 and A2), we use all five prior years as the reference years, rather than just

the 2018-19 school year. This approach allows us to present trend lines leading up to intervention

in the figure, but also allows us to leverage all pre-intervention data when reporting regression

coefficients. Our general conclusions are the same when we include all year dummy variables

and use the 2018-19 school year as the base year for comparison use.

Qualitative methods. We conducted one semi-structured hour-long interview with each of the instructional coaches during December 2020. Interviews were conducted over Zoom. The protocol elicited the coaches' perspectives on the implementation and outcomes of Ninth Grade Success in the pilot schools. Specific questions probed about who was involved with the team, how teams used data, and whether and to what extent schools saw improvements in student attendance, behavior, and course grades. We also asked about how the COVID pandemic affected the progress of Ninth Grade Success at the pilot schools. The interviews were transcribed by Zoom in real time.

To analyze the data, we started by reading and re-reading our notes from the interviews and the text of the transcripts. We then used a codebook to label and categorize the data. Codes were created inductively, based on patterns we were seeing across the transcripts (e.g., the importance of buy in), as well as deductively, based on key components of the model (e.g., data dashboards). This process allowed us to identify a few key themes that describe the pilot's implementation process and explain outcomes. To ensure trustworthiness, our initial write up was shared with the coaches and they were invited to make corrections and provide input. Coaches generally agreed with reported findings and did not request any substantial changes.

<div align="center">

Findings

</div>

In this section, we describe findings separately for our quantitative and qualitative analyses and then summarize our collective results.

Quantitative Findings

Results from the quantitative analysis are displayed in Figures 2 to 6. We start by presenting results from the Freshman Success Inventory survey, and then report findings for the ABC outcomes.

Freshmen Success Inventory survey. As shown in Figure 2, across all five pilot high schools and across all four domains, coaches reported increases in the appropriateness, level, or rigor of initial conditions, implementation, communication, and instruction. For example, at Ellensburg High School, across 13 items measuring the extent to which the principal, Team Lead, and Success Team set up conditions for success, the coach gave an average score of 2.08 for Fall 2019, corresponding roughly to "Emerging," whereas the same coach estimates an average score of 2.92 for Spring 2020, corresponding roughly to "Practicing." Ellensburg and Grandview High School tended to have larger scores in both the Fall and Spring; while Wahluke High School made the largest gains of any of the five schools in three of four categories (Moses Lake saw the highest gains for Setting Conditions). Moses Lake High School also started off at the lower level for all four domains, and typically finished the 2019-20 school year with the lowest scores. Results in Figure 2 suggest that all schools experienced gains in all four dimensions, but some schools made more progress than others. With that said, initial Fall score is a relatively strong predictor of Spring score, where schools starting lower tended to finish lower (Moses Lake and Wahluke), while those starting higher tended to finish higher (Ellensburg, Grandview, and Toppenish High Schools).

Attendance, behavior, and course grade outcomes. Next, we present results on the ABC outcomes in Figures 3 to 6, with additional results reported in the appendix. Panel A of Figure 3 shows the percent of ninth graders who attend at least 90% of school days each year. Pilot schools, represented by the solid black line, hovered around 70% of students from 2014-15 to 2019-20 and comparison group schools, marked in the gray dashed line, experienced a similar trajectory. This graph suggests that neither the treatment or control schools experienced significant changes in ninth grade regular attendance either before or after program

implementation. Figure 4, Panel A1 and A2 and Appendix Table A4 show that overall results for attendance are consistent for students classified as low-income or as English learners.

Panel B of Figure 3 shows results for the behavior outcome, which captures the severity of disciplinary incidence (increases in the behavior outcome represent decreases in discipline severity). Pilot high schools experienced a slight increase in the percent of disciplinary incidents resulting in fewer days of exclusion. In the years leading up to 2019-20, pilot and comparison group schools have similar behavioral outcomes. Pilot schools have a slightly higher rate of discipline severity compared to comparison schools in 2016 to 2018, but for 2019, the two groups both have almost the exact same percent of discipline incidence resulting in exclusion of three or fewer days (35%). In the first year of the program, pilot high schools see an increase in their behavioral outcome (i.e., a reduction in severity), though this increase is not statistically significant. Figure 4, Panel B1 and B2 and Appendix Table A4 show these results are driven by less severe discipline incidents specifically for students classified as low-income and for English Learners – and behavior outcomes for these two groups are statistically significant. Although not shown, we find null results for changes in overall discipline rates for treatment high schools (Figure 6, discussed below, shows overall discipline rates for individual treatment high schools).

Finally, Panel C of Figure 3 shows outcomes for course pass rates. In the years leading up to implementation, trends in course pass rates are generally parallel, but pilot schools have consistently lower pass rates than comparison group schools (although the difference is never statistically significant). In 2019-20, pilot schools experience a slight increase in course passing rates relative to comparison group schools, but the increase is small (about five percentage points) and not statistically significant. Appendix Table A3, which reports regression results using all pre-intervention years as reference years, shows treatment schools increase ninth grade

course pass rates during the treatment year by 2.6 percentage points more than comparison group schools, but again, this estimate is not statistically significant. Given the overall course pass rate of approximately 75 percent, a 2.6 percentage point increase may be considered educationally significant, particularly given the limited time frame that these estimated impacts represent. In Figure 4, Panel C1 and C2, and Appendix Table A4 show larger impacts for students classified as low-income or as English learners (but not statistically significant). Given the study's small sample size, including just five pilot high schools, we have limited power to detect small, but meaningful program impacts. Thus, while findings for course pass rates are not statistically significant, they provide suggestive evidence that the pilot caused increases in ninth grade course passing rates during the first year of implementation.

Given our finding of a slight improvement in behavior outcomes, Figures 5 and 6 provide additional insights into these results. First, Figure 5 shows changes in the severity of disciplinary incidents in the two years leading up to the intervention year and for the intervention year. For Toppenish High School, for example, about 15% of disciplinary incidents resulted in a student being excluded from school for more than 10 days during the 2017-18 school year. An additional 15% of disciplinary incidents resulted in a student being excluded from school for between 6 and 10 days. The following year, no students were excluded for more than 10 days, but in both years, over 40% of disciplinary incidents result in exclusion of four or more days. In contrast, in the first year of implementation, not one disciplinary incident at Toppenish High School results in exclusion of more than three days. Moses Lake High School and Grandview High School also experience decreases in the percent of disciplinary incidents that result in exclusion of more than three days, but Ellensburg and Wahluke both experienced increases in the severity of disciplinary incidents leading up to and including the intervention year.

Figure 6 helps put these findings in context. Both Toppenish and Grandview – the two schools showing declines in the severity of disciplinary incidents, also experience slight declines in 2019-20 in the overall percent of disciplinary incidents. Interestingly, Wahluke shows an even greater decline in disciplinary incidents, from about 9% in 2017-18 to 4% in 2019-20, the first year of the intervention, even though the severity increases. This may result from the waving of minor infractions, which would make punishments for "non-waved" disciplinary infractions appear more severe. When we consider as an outcome the total number of excluded days – a measure that captures both overall discipline rates and severity, our overall results are generally similar (a modest relative improvement for treatment schools during treatment years.

Qualitative Findings

Results of our qualitative analysis are reported in two parts, first examining factors that supported implementation and then assessing perceived positive outcomes of the pilot.

Factors that supported implementation.
"Because [the school] had the right makeup of [Success Team members], the right distribution of power, and the right resources directed to [the initiative], they could do it." —Center for High School Success Coach

The five schools in the pilot had varied experiences with implementing Ninth Grade Success Teams, with implications for long-term sustainability and student outcomes. The interviews with the coaches revealed several factors that supported implementation: (a) school buy in, (b) teacher leadership, (c) administrative support, and (d) local infrastructure. Each of these factors is discussed below.

School buy in. Per the Freshman On-Track model, a team should involve multiple stakeholders who work with freshmen, from administrators to teachers to support staff. One of the primary conditions that enabled implementation was buy in from the school community. Buy in ranged from interest among staff to serve on the team directly to widespread support for the

team's work. Coaches described well-functioning teams as those whose members were "big believers" in the model and "fully participatory." One coach stressed that collective buy in was perhaps the most important ingredient for success: "the greater the esprit de corps and an ethos and coherence around strategy… the greater the overall efficacy of the efforts for sure." Shared commitment to the initiative, in turn, motivated strategic adoption of the framework.

A crucial driver of buy in was relationship building. One coach's comment captured the essential nature of relationships for this work: "Relationships are important to get this really going." For successful implementation, relationship building was important at two levels. The first was among stakeholders in the school. The team lead played an important role in generating enthusiasm from his or her colleagues. For example, one lead was described as "incredibly organized and devoted," and "she got teachers on board right away." The second level of relationship building was between the coach and the team. Coaches recognized that "most school houses … get suspect of things that come in from the outside," so the first hurdle was "building trust with teams." The coaches cultivated those relationships by capitalizing on their years of principal experience, working collaboratively, helping teams see the model's potential to impact their school community, and encouraging them to take ownership of the process.

Teacher leadership. Teacher engagement and leadership were especially valuable to maximize Success Team implementation. "When… the teachers really own it," one coach explained, "it makes all the difference." Due to "how teachers view authority and credibility," they were more likely to embrace the work when the Freshman Success pilot did not seem like a top-down initiative. Thus, the coaches cautioned against an administrator leading the team. A coach described one Success Team where ninth grade teachers took the lead: "They had … the counselor and some support people but really [it was] the teachers. And… it took a little bit—but

[they] got comfortable with having a teacher lead the team. And that team as a whole was the most functional." The coaches acknowledged that shifting the distribution of power was challenging but had high pay off in terms of team members' investment in the process.

Teacher participation was also important for sustainability, because Freshman On-Track calls for shifting instructional practice. Coaches described a variety of team compositions, with some including fewer teachers than others. At one school, for example, no teachers wanted to be involved, so the team was comprised of counselors and student support staff. These staff worked hard to address students' immediate needs, such as offering an after-school program that served dinner. Coaches reported that these short-term interventions significantly improved the ninth grade on-track rate. However, referring to these afterschool programs, the coach explained, "if those folks from Gear UP and the support staff … stopped doing that, [the on-track rate would] drop right back down." In contrast, at schools where teachers were engaging in the work and embedding new practices into their classrooms, "a lot of the … growth will be attributed to more long-term change."

Administrative support. While coaches said administrators should not be team leads, they also stressed the need for administrator involvement and support. Coaches described principals, assistant principals, and in some cases district leaders being "very involved" and "supportive." One coach described how the principal and assistant superintendent "attended almost all the [team] meetings, which… showed very loud and clear their commitment." At another school, the principal was unable to attend the meetings but freed up the assistant principal to focus on the initiative. Participation from administrators was especially critical because they could make budgetary decisions that prioritized Ninth Grade Success. For example, when the Success Team at one school struggled due to a complicated master schedule with multiple lunch periods,

administration worked with the team to create a single ninth grade lunch, which reduced

opportunities for students to fall off track. As one coach put it, "whoever has the authority to

make decisions and parlay resources has to be right up in it with the team." Administrators took

on other larger school changes related to course programming and staff professional

development, and allocated funding and time to ensure these structures were in place.

Local infrastructure. Local infrastructure refers to specific structures in place that

facilitate the team's ability to meet during the regular school day and access and use the model

resources. The team's capacity to access and adapt particular resources to their local needs,

including time, data dashboards, and the Freshman-on-Track Toolkit was essential for supporting

implementation. The first important resource was time. For example, team members needed time

to attend professional development trainings and collaboratives where they could learn from and

with coaches and other schools in the Center for High School Success network. At the school

level, Success Teams needed shared meeting time at least once every other week. One coach

elaborated on the importance of these time allocations: "Meeting times have to be the same.

They have to be consistent. They have to be part of your week. They can't be something that is

thrown on [the schedule] at the last minute or changed." Success Teams that were not able to

gather on a regular basis and jointly analyze student data struggled to identify student needs and

develop interventions.

The data dashboard was another critical resource. Two issues with respect to data were

most salient: access and adaptation. Pilot schools had access to a data dashboard provided by

Center for High School Success, but access did not always translate to accessibility. Because

districts in Washington use local data information systems, a registrar or administrator at each

school had to be able to manipulate the student information system to make their ninth-grade

data easily viewable in the Freshman Success data dashboard. One school did not have this capacity. As a result, the team relied on one person to create usable data for each meeting, a process that not only taxed that person's time, but also undermined the team's ability to focus on students. At schools where the data were available in the dashboard, the teams "could literally just pull it up and project the data from homeroom on the screen for the team," which, a coach explained, "is how it should be." In addition to access, Success Teams needed to be able to adapt the data dashboard to capture the appropriate student groups. A coach recalled one school's experience with this particular hurdle: "There were some issues with [the dashboard] last year, fine tuning it… we needed it to show our English learners, our migrant students, our students with disabilities." Once the dashboard was updated to showcase those student groups, the team was "very happy with [it]." In short, infrastructure that allows team members to easily access and then adapt data to their student populations is critical for the work of the Success Team.

Finally, Success Teams needed to be able to modify resources from the Freshman On-Track Toolkit to account for the specific needs of their school community. The Toolkit includes the framework and numerous guidelines and protocols to support implementation, which the teams used "quite a bit," especially for communicating with parents and students. Yet as one coach pointed out, "You have to localize." At one school, "we probably have four different examples of a Kid Talk protocol," a protocol or script that guided the team's conversations about students and the corresponding interventions and helpful supports. Teams needed support to digest the resources provided in the Toolkit, determine what purpose they served and which were most applicable for their students, and adapt them accordingly.

Positive outcomes of the pilot.

"Whatever our numbers say right now, we are better off for having participated in this pilot."—Center for High School Success Coach

The coach interviews revealed several benefits of the pilot for participating schools, which are described across three themes: (a) positive stakeholder feedback, (b) cultural and structural shifts, and (c) improved capacity to respond to the COVID-19 pandemic. These insights contextualize and offer perspective not captured in the quantitative outcomes.

Positive stakeholder feedback. Broadly speaking, stakeholders who participated in the pilot saw value in the Success Team approach and were optimistic about its potential to improve student outcomes. As one coach said, "All schools saw this as something they came to rely on." At one school, the team found the model so useful that the district wanted to expand it down to eighth grade. Students, too, offered positive feedback. One coach relayed a comment the principal had shared with her: "The kids are coming to me, and they really want this for 10th grade. They liked it." Reflecting on why the students had this response, the coach posited, "See, they have those connections with staff members that were checking on them. It wasn't a punitive [check-in]. It was a, 'Let me come alongside you and support you.'" Such comments indicated that students and staff alike saw the model as a value add.

Cultural and structural shifts. The pilot schools also initiated cultural and structural change to varying degrees. These kinds of systemic shifts are slow-going and may not be reflected in short term quantitative indicators, but coaches expected the teams' investments would pay off in longer-term, sustainable change. Systemic change occurred in three main areas. The first pertained to the way in which stakeholders perceived their accountability to students. Reflecting on the team at one school, a coach described a "culture shift" from "individual teachers [seeing] their locus of control and responsibility as being in their classroom" to seeing "all of the kids that they taught as their responsibility to know what was going on for their … whole school experience." The model helped teachers broaden and reimagine the scope of their

accountability to their students beyond the confines of the 45 to 60 minutes students spent in their classroom.

The second shift involved more strategic thinking around how schoolwide structures and policies either supported or undermined student progress. For example, one school's team recognized that scheduling all freshmen to eat lunch together could promote their sense of belonging and community. Teams were also able to strategize around student programming and equitable parent outreach. Describing the Success Team at one school, a coach shared, "[They are] getting strategic about how to put students in what classes, particularly extra support classes, math support classes. Getting into the mechanics of making sure [there is language] accessibility… [for] a high Latino population."

Third, the pilot started to transform cultural beliefs around student evaluation. The coaches discussed a variety of "toxic grading practices" that increased students' likelihood of failure instead of assessing learning, such as assigning points to a parent permission slip or giving out zeros for missing assignments. Success teams were able to identify these practices when they analyzed student data. As one coach explained, "When … you can … look at [the student's] schedule of classes, and you see that they're failing advisory, that's an issue. You know, how do you fail advisory? … We were able to see [those patterns]." Advisory here refers to a class period in which students receive academic advising and catch up on schoolwork – a course that no student should be failing. The data allowed teams to identify these harmful practices in a "depersonalized" way—for example, to pinpoint whether the majority of failures were occurring in one course or one department without singling out individual teachers. From there, teams engaged in dialogues and professional development to understand equitable grading practices and evaluate whether their prior practices needed to change.

Improved capacity to respond to COVID-19. The COVID-19 pandemic presented a host of challenges for these schools beginning in March of the pilot year. Nationally, educators scrambled to keep track of students and ensure they could continue their coursework. One of the main obstacles for pilot schools was that students simply disappeared: "There were kids in these communities that just went missing," and many experienced mental health challenges such as anxiety and depression. The schools grappled with distributing devices and getting students on the internet. Once students had access, teachers struggled to keep them engaged online. Meanwhile, high schools had limited capacity to run school online, hybrid, and/or in person, or some combination of the three.

Fortunately, the ninth-grade teams were able to use their on-track infrastructure to bolster student support. Bi-weekly meetings to review student data, coupled with the personalized interventions that were already in place, allowed teams to better reach their students. As one coach elaborated, "the teams [could] really talk about individual kids and […] figure out where these kids were and what were the barriers … and get them the resources they needed." At one school, students' one-to-one mentors conducted home visits and "learned about the difficult situations students were facing that the school might not otherwise have been privy to." The coaches suspected that many more students would have become disconnected from schools and not received the assistance they needed had this pilot program not been in place.

Staff also reaped some benefits from Freshman Success during the pandemic. For example, meeting online instead of face-to-face meant that teams could connect more easily and frequently with their coaches. The coaches reported being able to attend more meetings and support more teams in a given day or week than they had prior to COVID. Virtual meetings also increased the teams' ability to access the knowledge and resources of other schools in the Center

for High School Success network: "We're able to provide professional development, pretty proficiently in the pandemic. I think we were good at creating the collaboratives pre-pandemic, but it required a lot of travel and a lot of extra logistical stuff. Now we can just sign in." In these ways, while COVID derailed the original plans for Ninth Grade Success implementation, teams and coaches adjusted and made the best of the circumstances of the pandemic.

Discussion

The Ninth Grade Success Initiative was implemented in five high schools beginning in the 2019-20 school year. While the COVID-19 pandemic changed many of the initial plans, the pilot served these schools well, providing opportunities to access and analyze data that, in some cases, better prepared educators for the pandemic's impact on schools and students. Our analysis demonstrates that coaches perceive the program positively, and implementation led to some small, but positive short-term effects, particularly around the severity of punishments related to disciplinary infractions and course pass rates. Impacts were larger for students classified as low-income and as English learners. We found no impacts on ninth grade student attendance. But these findings need to be placed in context, particularly the expectations one might hold for impacts from an intervention implemented over just five months, from the beginning of the school year in late September or early October 2019 to the end of February 2020. Given the typical effect size of educational interventions (e.g., Kraft, 2020), the effects on behavioral and course passing outcomes that we identify appear promising. Notably, two areas where we find the most change are those most under control of educators. While teachers and school staff are not able to force students to attend school, educators do have authority over how they conduct behavioral interventions and determine punishments and how they grade assignments and determine course grades.

These findings contribute to the literature in multiple ways. First, our study highlights an important role for coaches within early warning indicator interventions. Allensworth (2013), for example, argues that the freshman on-track indicator in Chicago Public Schools – which served as the basis for the Ninth Grade Success Initiative –helps educators focus conversations on actionable problems, identify students for intervention, and design strategic schoolwide reforms. All three of these positive outcomes are also outcomes associated with instructional coaching programs (Desimone & Pak, 2017; Knight, 2012; Kraft et al., 2018; Woulfin, 2018), but to our knowledge limited research examines the role of coaches within early warning systems.

Second, while most evaluations of early warning indicators are quantitative. (e.g., Bruce et al., 2011; Carl et al., 2013; Chung & Lee, 2019; Davis et al., 2019; Dynarski et al., 2008; MacIver et al., 2019), our use of multiple methods provided important insights. Coaches identified a number of strengths with the program that are likely to sustain as the program matures. In particular, coaches reported positive feedback from school leaders and a cultural shift in how teachers and other school staff approached their work. Key to the success of the initiative is community buy in, teacher leadership, administrative support, and local capacity..

Finally, little prior work examines the extent to which early warning or data-intensive interventions were useful during the COVID-19 pandemic (DeMatthews et al., 2020; Keller et al., 2020; Kwakye, 2021). On the one hand, such initiative may fall apart in the face of a major crisis. On the other hand, as much of our data showed, access to individual student data in an easy-to-use format can serve as an important resource during periods of crisis, allowing educators to quickly target their efforts to students or classrooms in need of tailored supports. In many ways, the pandemic created a stress test for the initiative, and educators demonstrated that

the major components of the initiative could be sustained through a major crisis and transition to virtual learning.

Findings for individual schools provide additional insights. As shown in Appendix Figure A1, Toppenish was the only high school to see overall increases in attendance during the intervention year, and that school produced increases both overall and for students classified as English learners and as low income. Four of the five schools reduced the severity of disciplinary incidents, both overall and for student groups, while Wahluke experienced increases in the severity of disciplinary infractions. Yet Wahluke High School had slight increases in course pass rates during the intervention year, along with Toppenish and Grandview. Toppenish High School was the only pilot high school to experience increases across all three outcomes for all students and student groups.

Based on these results, we argue this pilot program merits expansion to other high schools. We note that efforts to scale the program should ensure that the key elements of success are maintained. Moreover, we recommend that further inquiry into the success and challenges of the Pilot schools be conducted to help guide improvement and efforts to scale. Given this program's strong evidence-base and alignment with best practices (e.g., Roderick et al., 2011), our general findings that this program was successful are not surprising. Moving forward, state education leaders aiming to improve high school graduation rates should look to this pilot program as a promising initiative.

Conclusion

Aligning with the results of this evaluation, state education leaders and local districts in Washington view this intervention positively and have plans to maintain and expand the intervention moving forward. The districts involved in this study have all continued to

implement Ninth Grade Success Initiative. Through informal conversations, many have

expressed continued success with the program. Moreover, the state education agency has begun

facilitating expansion of the program beyond the initial set of five districts. Ongoing evaluation

will continue to explore experiences of teachers and other Success Team members as well as the

impacts on students.

The Ninth Grade Success Pilot was implemented in five high schools beginning in the

2019-20 school year. While the project was launched prior to the COVID-19 pandemic, our

qualitative data capture educator experiences through this period, and suggest that this pilot

program better prepared school leaders for the incoming crises and transition to remote learning.

By describing the implementation and outcomes of this college readiness reform, the study

informs broader efforts to improve educational systems and make them more responsive to
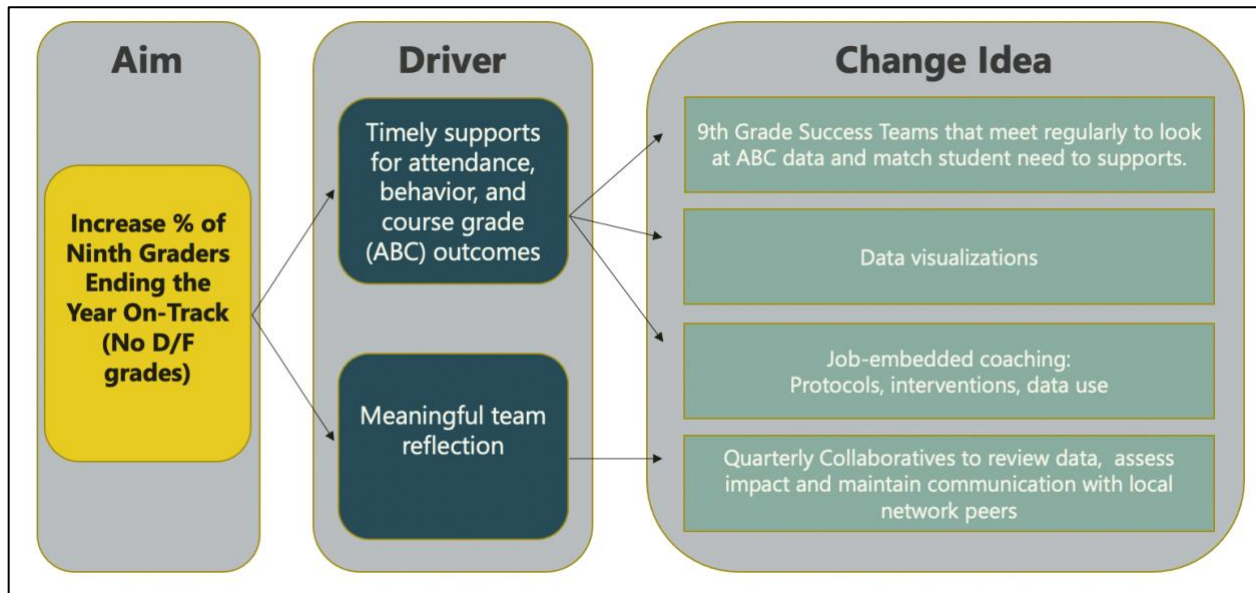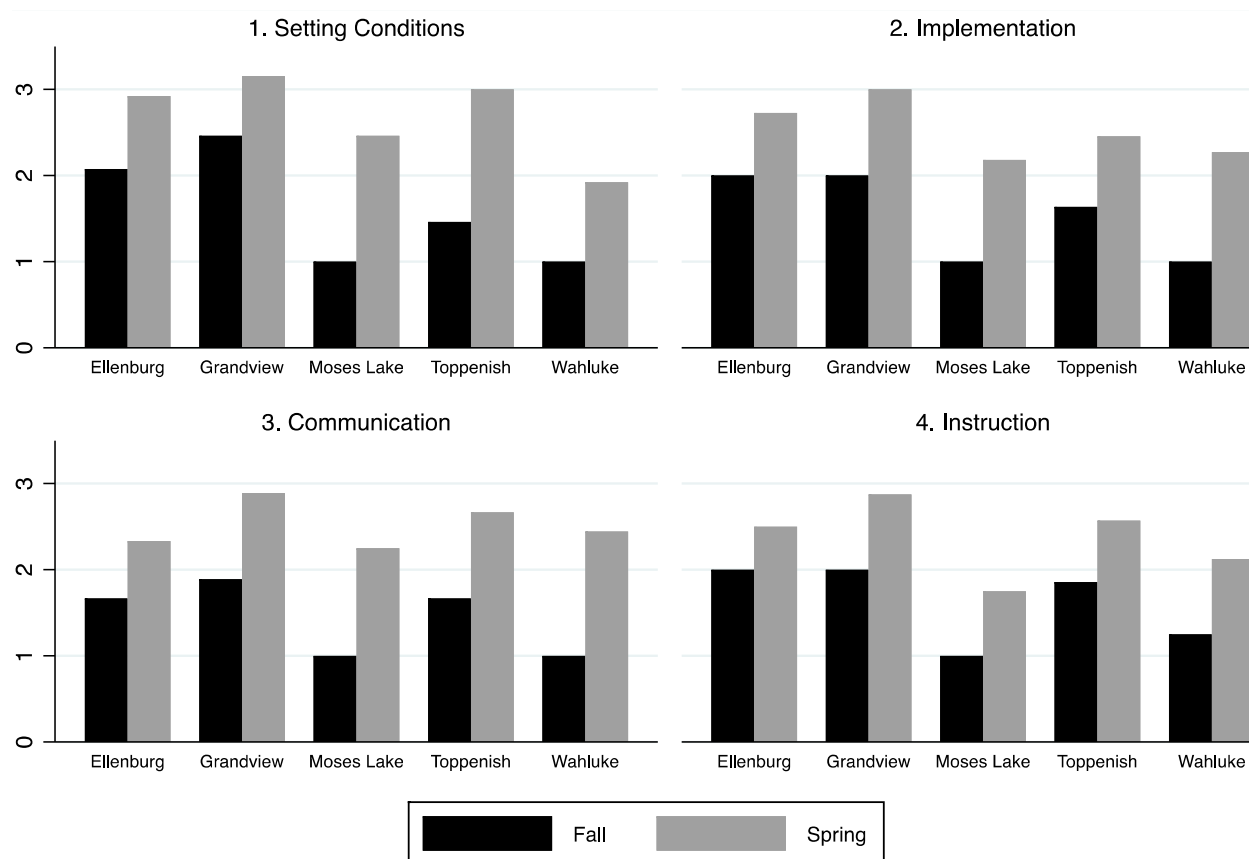
student needs.

FIGURE 1

Theory of change for the Ninth Grade Success Pilot

FIGURE 2

Survey results from the Freshman Success Inventory reflecting activities of the Principal, Team Lead, and Success Team across four domains, by pilot high school, 2019-2020


Note. Measures of setting conditions, implementation, communication, and instruction are based on 13, 11, 9, and 8 survey items, respectively, for which instructional coaches report on activities of the principal, Team Lead, and Success Team. For each survey item, coaches report on whether the activity reaches the following levels: 1= Not yet practicing; 2=Emerging; 3= Practicing; 4= Embedded. Thus, an average score of 2.0 for setting conditions (e.g., the Fall for Ellensburg High School) suggests that the appropriate conditions for the Freshman Success Framework are emerging, but not yet in practice. Appendix Table A1 displays each individual survey item and Appendix Table A2 shows raw survey results.

FIGURE 3

Regression-adjusted ninth grade student outcomes for pilot and comparison group schools, up to March 1 of each school year, 2014-15 to 2019-20

Panel A. Attendance
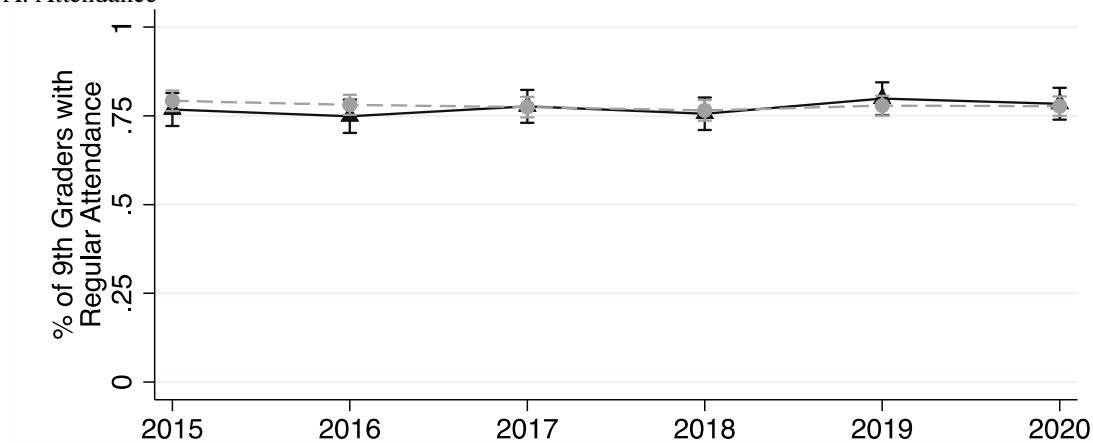
Panel B. Behavioral outcomes

Panel C. Course grades

Note. These figures plot marginal predictions based on regression coefficients from Equation 1. We use outcomes up to March 1 of each school to capture impacts during implementation year 1, 2019-20. Increases in the behavior outcome (Panel B) represent decreases in discipline severity.
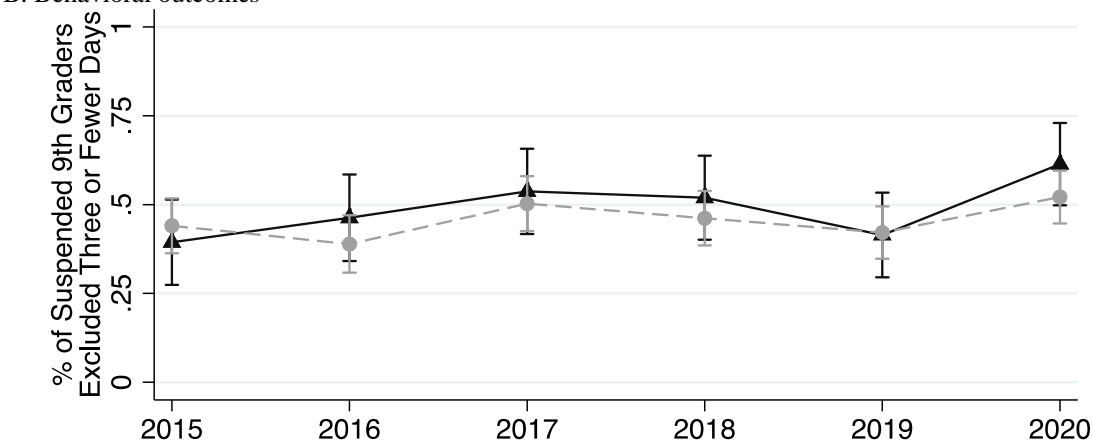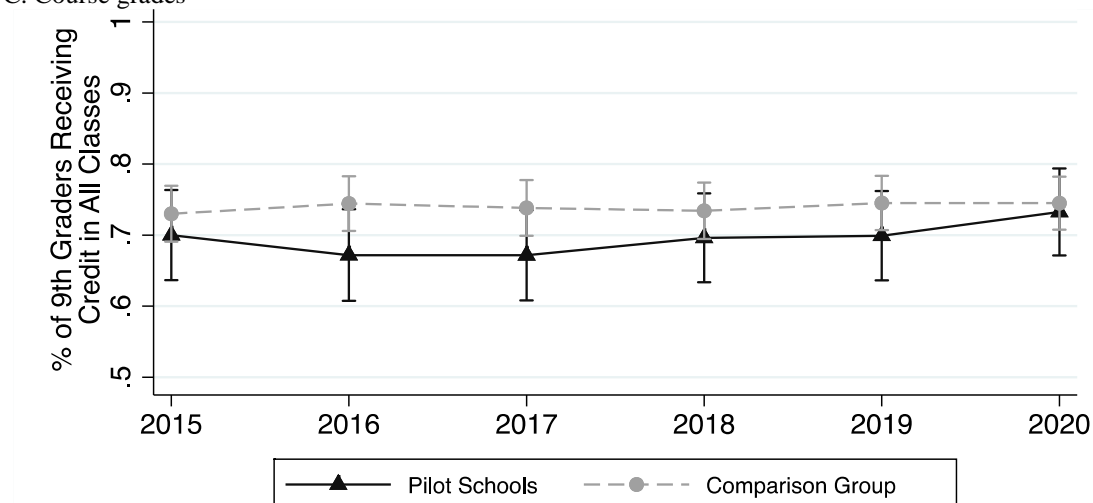
FIGURE 4

Regression-adjusted ninth grade student outcomes for pilot and comparison group schools, up to March 1 of each school year, 2014-15 to 2019-20

Panel A1. Attendance, students classified as FRL

Panel A2. Attendance, students classified as ELL
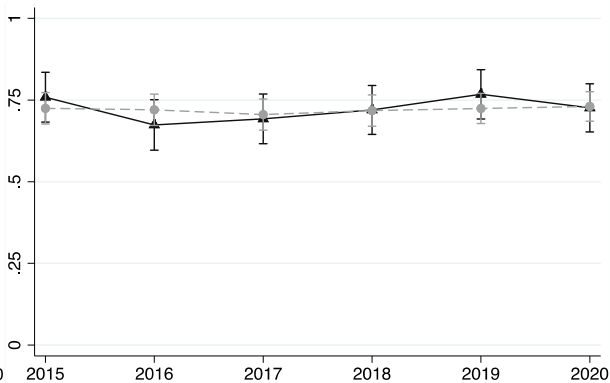
Panel B1. Behavior, students classified as FRL
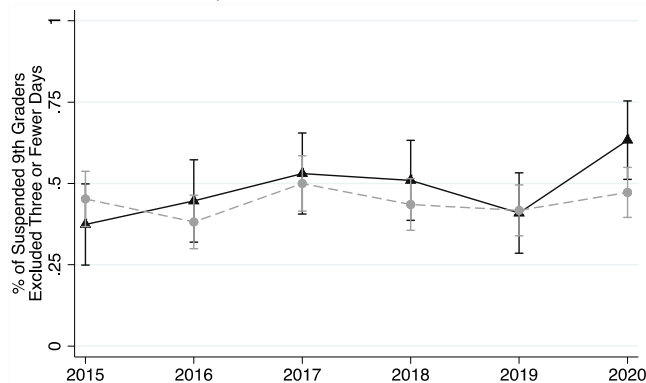
Panel B2. Behavior, students classified as ELL
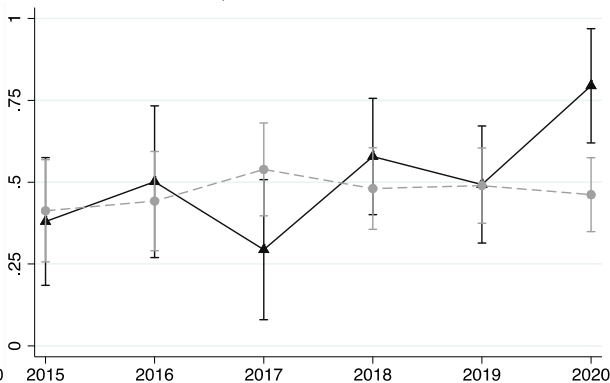
Panel C1. Course grades, stud. Classified as FRL
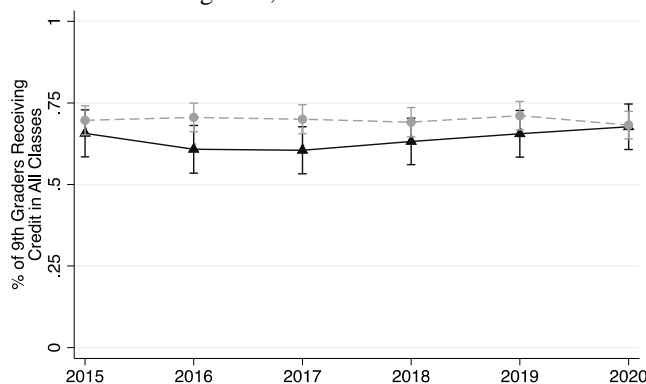
Panel C2. Course grades, stud. Classified as ELL
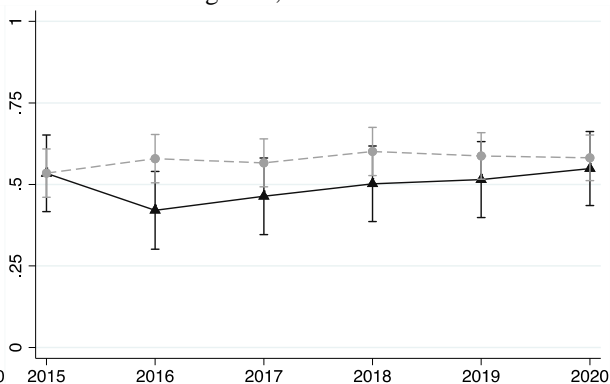
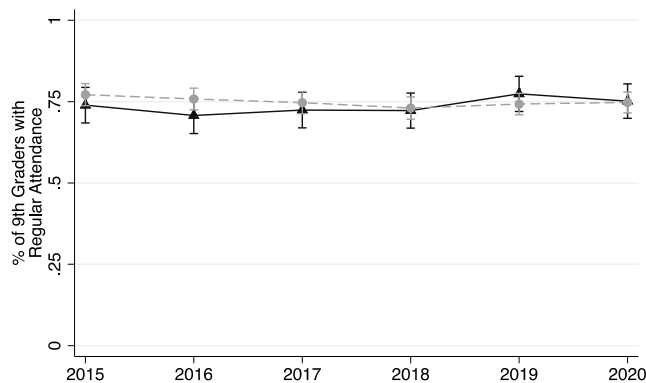*Note*. FRL=free/reduced price lunch, an indicator for students classified as low income. ELL=students classified as an English language learner. These figures plot marginal predictions based on regression coefficients from Equation 1. We use outcomes up to March 1 of each school to capture impacts during implementation year 1, 2019-20. Increases in the behavior outcome (Panels B1 and B2) represent decreases in discipline severity. Error bars reflect 90 percent confidence intervals. Regression coefficient for models that pool all pre-intervention years are shown in Appendix Table A3 and A4. *Source*: OSPI Report Card data, data as of February 26, 2021.

FIGURE 5

Percent of disciplinary incidents resulting in students excluded from class for disciplinary reasons, by number of days, pilot schools, 2017-18 to 2019-20
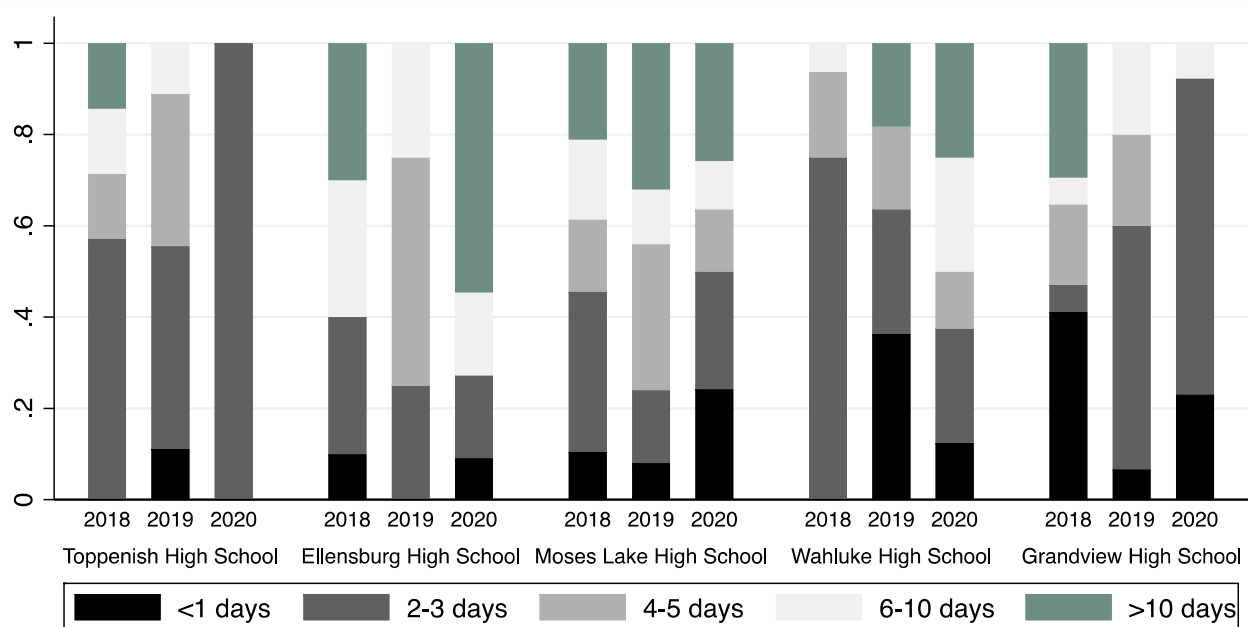
Source: OSPI Report Card data, data as of February 26, 2021.

FIGURE 6

Percent of ninth grade students excluded for disciplinary reasons for at least one day, treatment high schools, 2017-18 to 2019-20
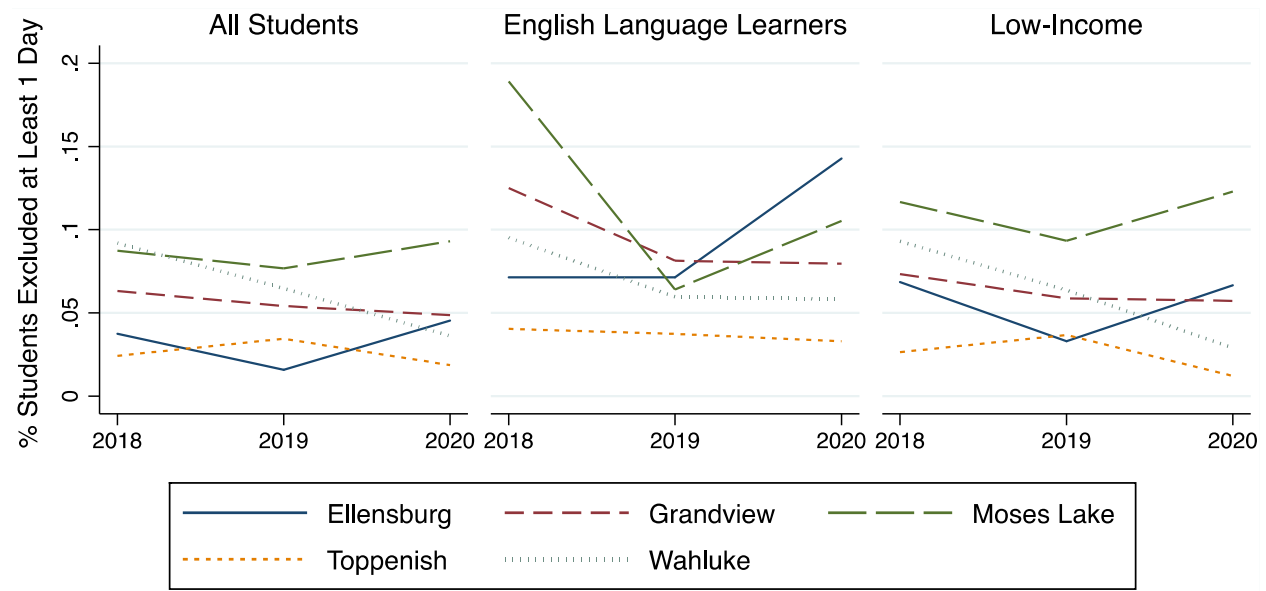
Source: OSPI Report Card data, data as of February 26, 2021. Data reflect outcomes up to March 1 of each year.

TABLE 1

Summary statistics for enrollment pilot schools, 2019-20

District / high school name	Enroll.	Student program			Student race/ethnicity						
		FRL	EL	SPED	Am.	As.	Bl.	His.	Pa.	Tm.	Wh.
Ellensburg	957	37%	5%	9%	1%	2%	1%	18%	0%	3%	75%
Grandview	1,009	72%	30%	13%	--	0%	--	93%	--	0%	7%
Moses Lake	2,420	60%	10%	14%	0%	1%	1%	47%	0%	3%	47%
Toppenish	948	88%	31%	15%	10%	--	0%	87%	--	1%	2%
Wahluke	670	53%	39%	17%	0%	--	--	97%	0%	--	3%

Note. Enrol.=Enrollment, FRL=Low-Income, EL=English learners, SPED=Students with disabilities, Am.=American Indian/Alaskan Native, As.=Asian Bl.=Black/African American, La.=Latinx/Hispanic, Pa.=Native Hawaiian/ Other Pacific Islander, Tm.= Two or More Races, and Wh.=White. Table reflects the whole school population, but enrollment characteristics are similar for ninth graders, the students targeted in the pilot. Source: OSPI Report Card data, data as of February 11, 2020.

TABLE 2

School characteristics by treatment group, 2018-19

	All high schools	Pilot high schools	Non-pilot HS in same county	Diff. from pilot HS	Non-pilot HS in other counties	Diff. from pilot HS
Schools	635	5	48	--	582	--
Enrollment	568	1,170	403	-767*	576	-594+
% Low-inc.	50%	71%	69%	-0.02	48%	-0.23*
% EL	8%	21%	17%	-0.04	7%	-0.14*
% SPED	17%	13%	13%	0.00	17%	0.04**
Amer. Ind.	4%	3%	9%	0.05	4%	0.01
Asian	6%	1%	1%	0.00	6%	0.05**
Black	6%	0%	1%	0.00	6%	0.05**
Latinx	22%	68%	50%	-0.17	19%	-0.48**
Pac. Is.	2%	0%	0%	0.00	2%	0.01
Two mo.	7%	2%	4%	0.02**	7%	0.06**
White	60%	27%	42%	0.15	61%	0.34*

Note. ** p<0.01, * p<.05, + p<0.10. Source: OSPI Report Card data, data as of February 11, 2020.

References

Allensworth, E. (2013). The use of ninth-grade early warning indicators to improve Chicago schools. *Journal of Education for Students Placed at Risk, 18*(1), 68-83.

Allensworth, E. M., & Easton, J. Q. (2007). What matters for staying on-track and graduating in Chicago public high schools: A close look as course grades, failure, and attendance in the freshman year. Chicago, IL: Consortium on Chicago School Research at the University of Chicago.

Arnold, K. D., Lu, E. C., & Armstrong, K. J. (2012). *The ecology of college readiness*. ASHE Higher Education Report Volume 38, Number 5. Wiley Periodicals, Inc.

Balfanz, R., & Byrnes, V. (2019). Early warning indicators and intervention systems: State of the field. In *Handbook of Student Engagement Interventions* (pp. 45-55). Academic Press.

Brown, T., Cook, A., & Santos, J. (2019). Reconceptualizing risk and high school noncompletion: The Case of Latina/o Ninth-Grade Leavers in an Urban School. Teachers College Record, 121(7), 1-32.

Bruce, M., Bridgeland, J. M., Fox, J. H., & Balfanz, R. (2011). On track for success: The use of early warning indicator and intervention systems to build a grad nation. Washington, DC: Civic Enterprises.

Butts, M. J., & Curxeiro, P. A. (2005). Student perceptions of factors leading to an effective transition from eighth to ninth grade. American Secondary Education, 42(1), 70–80.

Carl, B., Richardson, J. T., Cheng, E., Kim, H., & Meyer, R. H. (2013). Theory and application of early warning systems for high school and beyond. *Journal of Education for Students Placed at Risk, 18*(1), 29-49.

Chung, J. Y., & Lee, S. (2019). Dropout early warning systems for high school students using machine learning. Children and Youth Services Review, 96, 346-353. Cohen, J. S., & Smerdon, B. A. (2009). Tightening the dropout tourniquet: Easing the transition from middle to high school. Preventing School Failure, 53(3), 177–184.

Conley, D. T. (2008). Redefining college readiness. *New Directions for Higher Education, 144*, 3-13.

Davis, M., Herzog, L., & Legters, N. (2013). Organizing schools to address early warning indicators (EWIs): Common practices and challenges. Journal of Education for Students Placed at Risk (JESPAR), 18(1), 84– 100.

Davis, M., Mac Iver, M. A., Balfanz, R. W., Stein, M. L., & Fox, J. H. (2019). Implementation of an early warning indicator and intervention system. *Preventing School Failure: Alternative Education for Children and Youth, 63*(1), 77–88.

DeMatthews, D. E., Knight, D. S., Reyes, P., Benedict, A., & Callahan, R. (2020). From the field: Educational research during a pandemic. *Educational Researcher, 49*(6), 398-402, DOI: 10.3102/0013189X20938761

Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory into Practice, 56*(1), 3-12.

Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., & Smink, J. (2008). Dropout prevention: A practice guide (NCEE 2008-4025). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Fine, M., & Pryiomka, K. (2020). Assessing college readiness through authentic student work: How the City University of New York and the New York Performance Standards Consortium are collaborating toward equity. Palo Alto, CA: Learning Policy Institute.

Hoffman, J. L., & Lowitzki, K. E. (2005). Predicting college success with high school grades and test scores: Limitations for minority students. *The Review of Higher Education, 28*(4), 455-474.

Keller, M., DeBaun, B., & Warick, C. (2020). Supporting the Class of 2021 through postsecondary transition. Policy Brief. Washington, D.C.: Education Commission of the States.

Knight, D. S. (2012). Assessing the cost of instructional coaching. *Journal of Education Finance, 38*(1), 52-80.

Knight-Manuel, M. G., Marciano, J. E., Wilson, M., Jackson, I., Vernikoff, L., Zuckerman, K. G., & Watson, V. W. (2019). "It's all possible": Urban educators' perspectives on creating a culturally relevant, schoolwide, college-going culture for Black and Latino male students. Urban Education, 54(1), 35-64.

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

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547-588.

Mac Iver, M. A., Stein, M. L., Davis, M. H., Balfanz, R. W., & Fox, J. H. (2019). An efficacy study of a ninth-grade early warning indicator intervention. *Journal of Research on Educational Effectiveness, 12*(3), 363-390.

Mac Iver, M. A., Sheldon, S., Naeger, S., & Clark, E. (2017). Mentoring students back on-track to graduation: Program results from five communities. Education and Urban Society, 49(7), 643-675.

Martinez, M. A., Cortez, L. J., & Saenz, V. B. (2013). Latino parents' perceptions of the role of schools in college readiness. *Journal of Latinos and Education, 12*(2), 108-120.

McCallumore, K. M., & Sparapani, E. F. (2010). The importance of the ninth grade on high school graduation rates and student success. Education Digest, 130(3), 447–456.

McSpadden McNeil, L., Coppola, E., Radigan, J., & Vasquez Heilig, J. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. Education Policy Analysis Archives, 16(3), 1–44.

Nagaoka, J., Farrington, C. A., Roderick, M., Allensworth, E., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2013). Readiness for college: The role of noncognitive factors and context. *Voices in Urban Education, 38*, 45-52.

Neild, R. C. (2009). Falling off track during the transition to high school: What we know and what can be done. The Future of Children, 19(1), 53–76.

Office of Superintendent of Public Instruction (2020). Washington school improvement framework: Regular attendance. Retrieved from: https://www.k12.wa.us/sites/default/files/public/esea/essa/pubdocs/11wsifregularattendance.pdf.

Roderick, M., Nagaoka, J. Coca, V., & Moeller, E. (2008). From high school to the future: Potholes on the road to college. Chicago, IL: Consortium on Chicago School Research at The University of Chicago.

Roderick, M., Coca, V., & Nagaoka, J. (2011). Potholes on the road to college: High school effects in shaping urban students' participation in college application, four-year college enrollment, and college match. *Sociology of Education, 84*(3), 178-211.

Roybal, V., Thornton, B., & Usinger, J. (2014). Effective ninth-grade transition programs can promote success. Education, 134(4), 475–487.

Royster, P., Gross, J., & Hochbein, C. (2015). Timing is everything: Getting students back on track to college readiness in high school. The High School Journal, 98(3), 208-225.

Rumberger, R. W. (2011). Dropping out: Why students drop out of high school and what can be done about it. Cambridge, MA: Harvard University Press.

Venezia, A., & Jaeger, L. (2013). Transitions from high school to college. *The Future of Children*, 117-136.

Warren, J. R., Hoffman, E., & Andrew, M. (2014). Patterns and trends in grade retention rates in the United States, 1995-2010. Educational Researcher, 43(9), 433–443.

Welton, A., & Williams, M. (2015). Accountability strain, college readiness drain: Sociopolitical tensions involved in maintaining a college-going culture in a high" minority", high poverty, Texas high school. *The High School Journal, 98*(2), 181-204.

Wentworth, L., & Nagaoka, J. (2020). Chapter 1: Early warning indicators in education: innovations, uses, and optimal conditions for effectiveness. Teachers College Record, 122(14).

Woulfin, S. L. (2018). Mediating instructional reform: An examination of the relationship between district policy and instructional coaching. *AERA Open, 4*(3), 2332858418792278.

Appendix Tables

APPENDIX TABLE A1

Freshman Success Inventory survey items

---

1). SETTING CONDITIONS (13 items)

Principal
1   Selects, programs, and sets purpose around student success work for core set of grade-level teachers
2   Identifies team leadership and communicates expectations for engaging the Success Team and communicating progress
3   Sets goals for On-Track and student connection with the Success Team and team leadership
4   Provides and protects Success Team meeting calendar that includes regular and sufficient time for data analysis and intervention development, monitoring and adjustment
5   Provides time for General Education and Special Education teacher(s), counselor(s), student advocate or dean, and/or administration to attend regularly scheduled meetings

Team Lead
6   Acquires tools and strategies to lead the Success
7   With principal, establishes Success Team meeting calendar that includes regular and sufficient time for data analysis and intervention development, monitoring, and adjustment
8   With principal and Success Team, sets success goals for On- Track and student connection, and develops benchmarks to monitor progress

Success Team
9   Develops effective meeting strategies, such as building norms and action- oriented agendas
10   Shares Success Team roles and responsibilities conducive to success work
11   With principal and team leadership, sets purpose and student success goals for On- Track and student connection
12   Engages in regular calendared Success Team meetings to analyze data and develop, monitor, and adjust interventions
13   Includes General Education and Special Education teacher(s), counselor(s), student advocate or dean, and/ or administration

2). IMPLEMENTATION (11 items)

Principal
1   Provides timely access to success- related data such as point-in-time On-Track data
2   Provides professional development and training opportunities on high school transition
3   Reviews and interrogates interim success-related data in light of Success Team goals and strategizes with team leadership around next steps
4   Holds teachers accountable for implementing equitable grading practices

Team Lead
5   Develops action-oriented meeting agendas which consistently address success goals generally and intervention development, tracking, and evaluation specifically
6   Establishes Success Team meeting conditions conducive to the successful execution of Success Team duties
7   Works to bring actionable student- level data at regular intervals

Success Team
8   Reviews incoming class performance data to develop early and targeted supports for students
9   Develops, implements, tracks, and evaluates Level 2 interventions, making adjustments when appropriate
10   Refers students to appropriate level of Implementation
11   Reviews actionable student- level data in order to provide appropriate support

3). COMMUNICATION (9 items)

Principal
1   Provides space for student success-related communication on a regular basis
2   Explicitly supports Success Team efforts to communicate a culture of success for all

Team Lead

3   Regularly communicates strategies and progress toward success goals to Success Team, administration, parents, and students

4   Coordinates success assemblies, parent nights, and celebrations toward On-Track and student connection goals

5   Advocates for resources to support success efforts

Success Team

6   Maintains strengths-based and action- oriented communication to support students

7   Conducts joint parent/student conferences

8   Celebrates student and adult successes around On-Track and student connection goals

9   Engages faculty in frequent communication on student progress and successful strategies

4). INSTRUCTION (8 items)

Principal

1   Communicates a clear vision for instruction

2   Supports teacher development through proactive observation and constructive feedback around instructional practice

3   Provides learning opportunities on classroom practices that support students' successful transition to high school

Team Lead

4   Seeks out instructional and intervention resources, readings, and tools to improve teacher and Success Team practice

5   Designs and facilitates Success Team discussion, problem solving, and sharing around grading and instructional strategies

Success Team

6   Creates, implements, and evaluates instructional strategies around student engagement

7   Provides multiple and varied opportunities for students to exhibit mastery or be assessed

8   Utilizes transparent and equitable grading practices that communicate grade-level expectations and student achievement

APPENDIX TABLE A2

Freshman Success Inventory survey results

	EHS		GHS		MLHS		THS		WHS	
	F	S	F	S	F	S	F	S	F	S

1). SETTING CONDITIONS (13 items)

Principal

		EHS F	EHS S	GHS F	GHS S	MLHS F	MLHS S	THS F	THS S	WHS F	WHS S
1	Selects, programs, and sets purpose around student success work for core set of grade-level teachers	3	3	3	4	1	2	2	3	1	2
2	Identifies team leadership and communicates expectations for engaging the Success Team and communicating progress	2	2	3	4	1	3	2	3	1	1
3	Sets goals for On-Track and student connection with the Success Team and team leadership	2	3	3	3	1	2	2	3	1	3
4	Provides and protects Success Team meeting calendar that includes regular and sufficient time for data analysis and intervention development, monitoring and adjustment	2	3	3	3	1	3	1	3	1	2
5	Provides time for General Education and Special Education teacher(s), counselor(s), student advocate or dean, and/or administration to attend regularly scheduled meetings	2	3	3	3	1	1	1	3	1	2

Team Lead

6	Acquires tools and strategies to lead the Success	2	3	2	3	1	3	2	3	1	2
7	With principal, establishes Success Team meeting calendar that includes regular and sufficient time for data analysis and intervention development, monitoring, and adjustment	2	3	2	3	1	3	2	3	1	2
8	With principal and Success Team, sets success goals for On-Track and student connection, and develops benchmarks to monitor progress	2	3	3	3	1	3	2	3	1	2

Success Team

9	Develops effective meeting strategies, such as building norms and action- oriented agendas	2	3	2	3	1	3	1	3	1	2
10	Shares Success Team roles and responsibilities conducive to success work	2	3	2	3	1	2	1	3	1	2
11	With principal and team leadership, sets purpose and student success goals for On- Track and student connection	2	3	2	3	1	2	1	3	1	2
12	Engages in regular calendared Success Team meetings to analyze data and develop, monitor, and adjust interventions	2	3	2	3	1	3	1	3	1	2
13	Includes General Education and Special Education teacher(s), counselor(s), student advocate or dean, and/ or administration	2	3	2	3	1	2	1	3	1	1

2). IMPLEMENTATION (11 items)

Principal

1	Provides timely access to success- related data such as point-in-time On-Track data	2	3	-	-	1	3	2	3	1	2
2	Provides professional development and training opportunities on high school transition	2	3	-	-	1	2	1	2	1	2
3	Reviews and interrogates interim success-related data in light of Success Team goals and strategizes with team leadership around next steps	2	3	-	-	1	3	2	3	1	3
4	Holds teachers accountable for implementing equitable grading practices	2	3	-	-	1	1	1	2	1	2

Team Lead

5	Develops action-oriented meeting agendas which consistently address success goals generally and intervention development, tracking, and evaluation specifically	2	3	-	-	1	3	2	3	1	3
6	Establishes Success Team meeting conditions conducive to the successful execution of Success Team duties	2	3	-	-	1	2	2	3	1	2
7	Works to bring actionable student- level data at regular intervals	2	3	-	-	1	2	2	3	1	3

Success Team

8	Reviews incoming class performance data to develop early and targeted supports for students	2	2	-	-	1	2	2	2	1	2
9	Develops, implements, tracks, and evaluates Level 2 interventions, making adjustments when appropriate	2	2	-	-	1	2	1	2	1	2
10	Refers students to appropriate level of Implementation	2	2	-	-	1	2	2	2	1	2
11	Reviews actionable student- level data in order to provide appropriate support	2	3	2	3	1	2	1	2	1	2

3). COMMUNICATION (9 items)

Principal

| 1 | Provides space for student success-related communication on a regular basis | 2 | 3 | 3 | 4 | 1 | 2 | 1 | 2 | 1 | 2 |
| 2 | Explicitly supports Success Team efforts to communicate a culture of success for all | 2 | 3 | 3 | 3 | 1 | 2 | 2 | 3 | 1 | 3 |

Team Lead

3	Regularly communicates strategies and progress toward success goals to Success Team, administration, parents, and students	2	2	2	3	1	2	2	3	1	2
4	Coordinates success assemblies, parent nights, and celebrations toward On-Track and student connection goals	2	3	1	3	1	3	3	3	1	3
5	Advocates for resources to support success efforts	1	1	1	2	1	2	1	3	1	3

Success Team

6	Maintains strengths-based and action- oriented communication to support students	2	3	2	3	1	2	2	3	1	2
7	Conducts joint parent/student conferences	2	2	3	3	-	-	2	2	1	2
8	Celebrates student and adult successes around On-Track and student connection goals	1	2	1	2	1	2	1	3	1	3
9	Engages faculty in frequent communication on student progress and successful strategies	1	2	1	3	1	3	1	2	1	2

4). INSTRUCTION (8 items)

Principal

1	Communicates a clear vision for instruction	3	3	3	3	1	2	2	3	1	2
2	Supports teacher development through proactive observation and constructive feedback around instructional practice	3	3	3	3	1	2	-	-	3	3
3	Provides learning opportunities on classroom practices that support students' successful transition to high school	2	2	3	3	1	2	3	3	1	2

Team Lead

| 4 | Seeks out instructional and intervention resources, readings, and tools to improve teacher and Success Team practice | 2 | 3 | 2 | 3 | 1 | 2 | 3 | 3 | 1 | 2 |
| 5 | Designs and facilitates Success Team discussion, problem solving, and sharing around grading and instructional strategies | 2 | 3 | 1 | 3 | 1 | 2 | 2 | 3 | 1 | 2 |

Success Team

6	Creates, implements, and evaluates instructional strategies around student engagement	1	2	1	3	1	2	1	2	1	2
7	Provides multiple and varied opportunities for students to exhibit mastery or be assessed	2	2	1	3	1	1	1	2	1	2
8	Utilizes transparent and equitable grading practices that communicate grade-level expectations and student achievement	1	2	2	2	1	1	1	2	1	2

APPENDIX TABLE A3

Difference-in-difference regression coefficients assessing the impact of Ninth Grade Success Pilot on student attendance, behavioral outcomes, and course pass rates up to March 1 of each school year, 2014-15 to 2019-20

	Attendance	Behavior	Course pass rates
Treatment schools x treatment year	0.013	0.076	0.026
	(0.038)	(0.096)	(0.046)
Treatment schools	0.014	0.043	-0.012
	(0.018)	(0.047)	(0.022)
Treatment year (2019-20)	-0.007	0.079+	-0.001
	(0.020)	(0.053)	(0.025)
Covariates			
Log enrollment	-0.031***	-0.057**	-0.037***
	(0.009)	(0.025)	(0.011)
% free/reduced price lunch	-0.182***	0.034	-0.397***
	(0.064)	(0.180)	(0.078)
% students rec. SPED services	-0.104	-0.281	0.219
	(0.195)	(0.544)	(0.239)
% student classified as ELL	0.067	0.135	-0.089
	(0.084)	(0.225)	(0.104)
Constant	1.080***	0.746***	1.216***
	(0.068)	(0.206)	(0.083)
Num. of high school-year obs.	169	143	169
R-squ.	0.168	0.102	0.368

Note. Row 1 reports coefficients for the interaction term for the treatment school indicator and the treatment year, which provides the causal impact of the Ninth Grade Success Initiative on the three outcomes of interest, attendance, behavior, and course pass rates, defined in the text. We focus on outcomes up to March 1 of each school year so that outcomes for the treatment year, 2019-20 are assessed prior to COVID-19 school closures. All models include county fixed effects and the sample is limited to treatment high schools and all other high schools located in one of the three counties that contain at least one treatment high school (n=169 high school-year observations). A total of 27 high school-year observations are missing behavior data due to small cell size.
+ $p<0.20$, * $p<0.10$, ** $p<.05$, *** $p<0.01$.

APPENDIX TABLE A4

Difference-in-difference regression coefficients assessing the impact of the Ninth Grade Success Pilot on student attendance, behavioral outcomes, and course pass rates up to March 1 of each school year, 2014-15 to 2019-20

	Attendance		Behavior		Course pass rates	
	FRL	ELL	FRL	ELL	FRL	ELL
Treatment schools x treatment year	0.024	-0.013	0.155+	0.391***	0.060	0.029
	(0.044)	(0.058)	(0.098)	(0.137)	(0.055)	(0.083)
Treatment schools	0.006	0.02	0.037	0.019	-0.026	-0.036
	(0.021)	(0.028)	(0.050)	(0.075)	(0.027)	(0.040)
Treatment year (2019-20)	-0.009	-0.003	0.042	-0.017	-0.029	-0.011
	(0.024)	(0.032)	(0.055)	(0.079)	(0.030)	(0.045)
Covariates						
Log enrollment	-0.040***	-0.023*	-0.062**	-0.104**	-0.054***	-0.070***
	(0.011)	(0.014)	(0.026)	(0.042)	(0.013)	(0.020)
% free/reduced price lunch	-0.091	-0.190*	0.142	-0.039	-0.315***	-0.506***
	(0.074)	(0.098)	(0.188)	(0.304)	(0.093)	(0.141)
% students rec. SPED services	-0.26	0.158	-0.385	-0.782	0.139	0.762*
	(0.228)	(0.301)	(0.577)	(0.893)	(0.286)	(0.436)
% student classified as ELL	0.156+	0.147	0.082	-0.386	-0.024	0.093
	(0.099)	(0.130)	(0.232)	(0.371)	(0.123)	(0.187)
Constant	1.038***	0.940***	0.713***	1.275***	1.212***	1.221***
	(0.080)	(0.105)	(0.217)	(0.349)	(0.100)	(0.151)
N	169	168	134	96	168	165
R-squ.	0.172	0.111	0.117	0.213	0.233	0.217

Note. Row 1 reports coefficients for the interaction term for the treatment school indicator and the treatment year, which provides the causal impact of the Ninth Grade Success Initiative on the three outcomes of interest, attendance, behavior, and course pass rates, defined in the text. We focus on outcomes up to March 1 of each school year so that outcomes for the treatment year, 2019-20 are assessed prior to COVID-19 school closures. All models include county fixed effects and the sample is limited to high schools located in one of the three counties that contain at least one treatment high school. Changes to the sample are due to missing values for some comparison group high schools.
+ $p<0.20$, * $p<0.10$, ** $p<.05$, *** $p<0.01$.