Success in the college preparatory mathematics pipeline: Impact of policies and practices employed by three high school reform models

By: Nina Arshavsky, Julie A. Edmunds, Luke C. Miller, and Matthew Corritore.

Arshavsky, N., Edmunds, J.A., Miller, L., & Corritore, M. (2014). Success in the college preparatory mathematics pipeline: Impact of policies and practices employed by three high school reform models. *School Effectiveness and School Improvement*: *An International Journal of Research, Policy and Practice, 25*:4, 531-554.

This is an Accepted Manuscript of an article published by Taylor & Francis in *School Effectiveness and School Improvement*: *An International Journal of Research, Policy and Practice* on 10/29/2013, available online: http://dx.doi.org/10.1080/09243453.2013.846917

Abstract:

This paper examines the relationship of the policies and practices employed by 3 high school reform models – Early College High Schools, Redesigned High Schools, and High Schools That Work – with student success in college preparatory mathematics courses by the end of the 10th grade. Data on policies and practices collected through a survey of school principals in North Carolina are combined with administrative data on student course-taking and performance. The examined policies include course-taking requirements, rigorous instruction, academic support, personalization, and relevance. Results show that implementation of these policies varies across models and that higher levels of implementation of combinations of these policies are associated with improved outcomes.

Keywords: high school reform | mathematics | college readiness

Article:

Introduction

To respond to recent concerns about high levels of post-secondary remediation and insufficient workforce preparation (Achieve, Inc., 2004), national organizations, states, and districts have begun efforts to ensure more students graduate prepared for college and the workplace. Given research showing a rigorous high school curriculum, particularly in mathematics, is the best indicator of success in college (Adelman, 2006), many states and districts have begun to change their graduation requirements to establish a default expectation that all students take a sequence of college preparatory mathematics courses (Edmunds & McColskey, 2007).

Although a fair amount of research has already been done on course-taking policies, (Finn, Gerber, & Wang, 2002; Gamoran & Hannigan, 2000; Lee, Croninger, & Smith, 1997; Lee & Smith, 1999), there is less research on the policies and practices that support students in taking and succeeding in these more rigorous courses. “Researchers and practitioners must examine the processes within schools that lead to some students taking advanced coursework but not others…” (Finn et al., 2002, p. 365). This article contributes to the research base by examining school-level policies and practices that are designed to improve mathematics course-taking within the context of comprehensive school reform.

Analyzing data on three high school reform models common in North Carolina – Early College High Schools (ECHS), Redesigned High Schools (RHS), and High Schools That Work (HSTW) – this study examines the implementation of a core set of school-level policies and practices and its association with student academic performance in college-preparatory mathematics courses. Results show that implementation of these policies varies across models. Additionally, higher levels of implementation of combinations of these policies are associated with improved student outcomes, while increases in individual policies show a weaker association with outcomes.

Theoretical background

This paper is part of a larger study examining the impact of three high school reform models on students’ college preparatory mathematics and science course-taking and success. We examine three reform models widely implemented throughout North Carolina that aim to ensure all students, including students underrepresented in college populations such as low-income, minority, and first-generation college-goers, successfully progress through a college preparatory curriculum. In addition to looking at the overall impact of each model in a larger study, this paper examines the specific policies and practices in each model that are intended to influence students’ course-taking and success. Figure 1 shows the core components of the reform models, specific policies and practices under each core component, and the expected outcomes. This figure guides the design of our study.

Figure 1. Theory of change, high school reform models.

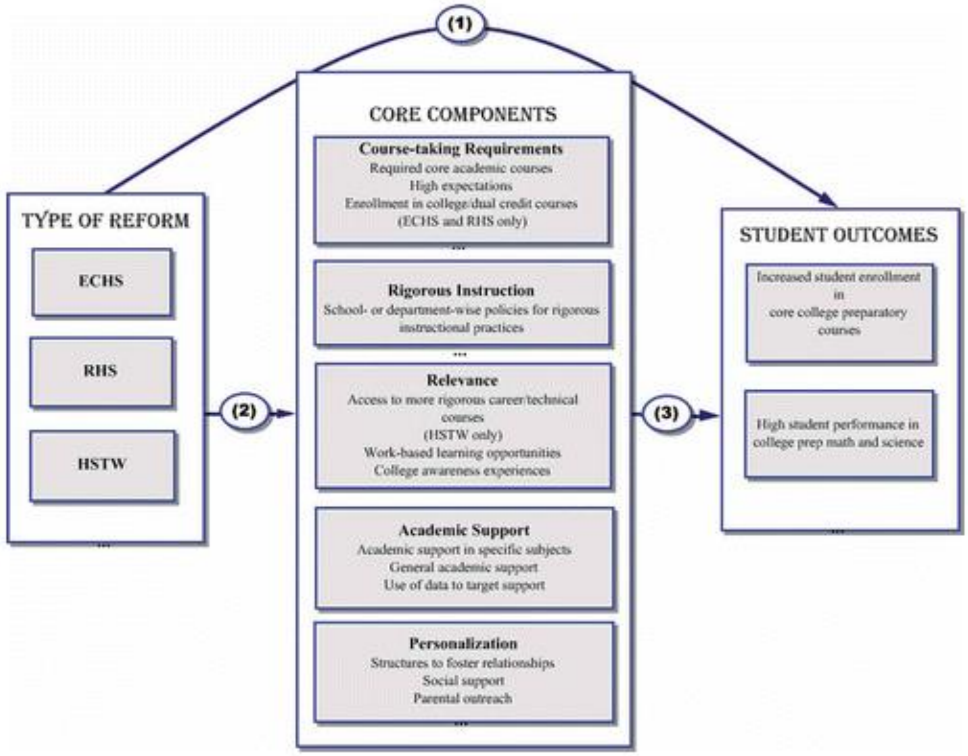


Figure 1 also shows the relationships between the policies and practices and outcomes that we tested. The first relationship (shown as Path 1) is the direct impact of the entire model on student outcomes; this relationship is examined in other papers (Miller & Corritore, 2011; Miller & Mittleman, 2011, 2012). The second relationship (Path 2) is the association between the type of reform model and the level of implementation of these different categories of policies and practices. The final relationship (Path 3) is the association between the implementation of these policies and practices and student outcomes. This paper focuses on the second and third relationships: how well schools adopting one of the three selected reform models implement the five categories of policies and practices, and how implementation of these policies and practices is associated with student learning outcomes.

Numerous studies on and evaluations of reform efforts demonstrate that implementation fidelity varies significantly among and within reform models and depends on such additional factors as buy-in of school staff and support from the district and parents among others (Berends, Bodilly, & Kirby, 2002; Berman & McLaughlin, 1978; Kirby, Berends, & Naftel, 2001; Supovitz &Weinbaum, 2008; Weinbaum & Supovitz, 2010). From our other studies on implementation of ECHS in North Carolina, we have indications that the buy-in among ECHS staff is often strong, and sometimes stronger than the buy-in of staff in Redesign schools (unpublished data and personal communications with the North Carolina New Schools staff). Based on these data and research, we predicted that we may see a stronger implementation of selected policies in ECHS schools than in schools of two other types.

Additionally, previous studies have shown that with reform models using effective strategies, stronger model implementations lead to stronger student achievement results (Berman & McLaughlin, 1978; Datnow, Borman, & Stringfield,2000; Datnow, Lasky, Stringfield, & Teddlie, 2005). Based on this research and research reviewed below (about the effectiveness of policies), we predicted that we will see positive associations between higher levels of policies implementation and student success in the mathematics pipeline.

We begin by briefly describing the three models, and then describe the research on policies and practices examined in the present study that all three models have in common.

High school reform models

The three models examined are Early College High Schools, Redesigned High Schools, and High Schools That Work. Early College High Schools (ECHS) are small schools, most often located on college campuses, designed to blur the distinction between high school and college. Serving students in Grades 9 to 12 (with some also serving Grade 13), the ECHS model is targeted at students who are underrepresented in college, including students who are low-income, the first in their family to go to college, or members of underrepresented ethnic and racial groups. All ECHS schools in this sample were newly created within 4 to5 years prior to this study. Both faculty and students had to apply to this type of school (rather than to be assigned), which could differentiate this type of schools in some unmeasured ways. Students attending an ECHS are expected to graduate in 4 to 5 years with a high school diploma and up to 2 years of college credit. Redesigned High Schools (RHS) are small schools carved out of a large, comprehensive high school and are organized around themes. An RHS may be a single school spun off from a larger high school, or may be part of a large school that has been completely broken up into a set of independent small high schools all on the same campus. Part of the theory underlying these schools is that students and, to a lesser degree, faculty are able to choose a school with a theme that is most aligned to their interests. However, in situations where an entire school devolved into small schools, all students and faculty needed to be placed in one of these schools. As a result, students and faculty may or may not have applied to individual RHS in this sample.

In North Carolina, both RHS and ECHS are guided by the same six Design Principles developed by the North Carolina New Schools, the non-profit organization overseeing the implementation of these models (North Carolina New Schools, n.d.). These design principles are: a focus on college readiness, including a default college preparatory curriculum and access to college credit courses; teaching and learning that emphasize critical thinking, application, and problem-solving skills; a personalized learning environment with strong staff–student relationships and affective and academic support; a professional working climate where teachers collaborate and collectively take responsibility for student learning; leadership that works to develop a common vision; and purposeful design that supports the other principles, including a small size of less than 400 students and flexible use of time. For both of these models, students had to actively select into the school.

High Schools That Work (HSTW) is a comprehensive high school reform effort working to create a “culture of high expectations that motivates students to make the effort to succeed in school” (Southern Regional Education Board [SREB],n.d.). Developed by the SREB, the HSTW model incorporates 10 *Key Practices* designed to achieve this goal. These practices include high expectations, a rigorous academic college-preparatory curriculum combined with career and technical concentration, work-based learning, teacher collaboration, rigorous instructional strategies, a guidance and advisory system, additional academic supports, and a culture of continuous improvement throughout the school. These practices are designed to “help all students leave high school with an employer certification, postsecondary credit, or the knowledge and skills needed to avoid remedial postsecondary studies” (SREB, n.d.).

Policies and practices

The brief descriptions of each model highlight their common goal of increasing the number of students who graduate ready for postsecondary education. Within each of these models, there are research-based policies and practices designed to increase the number of students who are succeeding in college preparatory courses. We identified five categories of policies and practices that are common among the three reform models and that have been shown in prior research to improve outcomes associated with increased enrollment and success in college preparatory mathematics and science courses (Balfanz, McPartland, & Shaw, 2002; Finn et al., 2002; Lee et al., 1997; Mayer, 2008; Secada et al., 1998). These policies and practices are shown in Figure 1.

Course-taking requirements

All three models propose to enroll more students in a college preparatory course of study. Research has shown that a challenging high school curriculum is associated with higher student achievement (National Center for Education Statistics [NCES], 2004) and with success in college, even among students with certain risk factors such as being a first-generation college student (Adelman, 1999; NCES, 2001). Lee and Burkam (2003) found schools with more academic and fewer remedial courses also had higher graduation rates regardless of students’ academic background and school performance. However, examinations of recent efforts to mandate core college preparatory courses for all students, such as Algebra I by the end of ninth grade, have found that simply increasing course-taking requirements may increase the number of students who take these courses but not necessarily improve their mathematics achievement (Allensworth, Nomi, Montgomery, & Lee, 2009). Other studies have suggested increasing course-taking requirements may also increase dropout rates (Jacob, 2001; Lillard & DeCicca, 2001).

Rigorous instruction

All three models attempt to increase the rigor of classroom instruction. Previous studies have shown that improved student achievement is linked to more rigorous student assignments, including asking students to engage in higher level thinking, apply core content, and provide extended explanations of information (Newmann, Bryk, & Nagaoka, 2001; Newmann, Lopez, & Bryk, 1998). Studies have found specific positive relationships between the rigor of instruction in mathematics classrooms and the quality of student work and subsequent test scores (Mitchell et al., 2005; Shkolnik et al., 2007). Additionally, positive relationships have been found between rigorous instructional strategies, such as engagement in higher order thinking, and the teachers' value-added scores (MET Project, 2010, 2012).

Relevance

These high school reform models emphasize the relevance of high school education by connecting it to students' lives outside of the classroom and beyond high school. Such connections may involve a theme-based course of studies as in RHS, work-based learning and blended academic and career content of studies as in HSTW, and advising for college as in RHS and ECHS.

Evaluations of these reform models show that theme-based learning communities can play a role in increasing attendance and reducing dropout rates, although not necessarily in improving students' test scores (Kemple, 2008). Coordinating technical and academic courses can improve students' math skills (Stone, Alfeld, & Pearson, 2008), although work-based learning has been found to have a potentially negative impact on students' classroom performance (Kaufman, Bradby, & Teitelbaum, 2000; Stasz & Brewer, 1998).

Relevance can also be seen as making a connection between high school and college. In an experimental study conducted on ECHS by two of the authors of this paper, the researchers found that the ECHS' strong link to college increased the expectations for students and made them want to perform better in their classes (Edmunds, Willse, Arshavsky, & Dallas,2013). The opportunity to take college credit courses has also been associated with increased graduation rates among students (Karp, Calcagno, Hughes, Jeong, & Bailey, 2007).

Academic support

Studies suggest that in order for a wider range of students to succeed in more challenging courses, they need additional academic support (Mayer, 2008; Swanson, Mehan, & Hubbard, 1995). Recent evaluation studies of high school reforms also report on the positive influence of academic support, as part of the reform design, on student academic and behavioral outcomes (Bloom, Thompson, & Unterman, 2010; Stevens, Sporte, Stoelinga, & Bolz, 2008). Associated with improved student academic outcomes are different types of academic support including peer tutoring (Robinson, Schofield, & Steers-Wentzell,2005; Spencer, 2006), volunteer tutoring (Ritter, Barnett, Denny, & Albin, 2009), summer and after-school programs (Lauer et al., 2006), and extended classes that serve as a "double-dose" of the subject matter (Takoko & Allensworth, 2009).

Personalization

All three models attempt to personalize the educational experience through guidance and advising (HSTW) or through positive relationships and affective support (ECHS and RHS). Studies show that when students feel that they are noticed, known, and cared for by adults, they are more engaged, more motivated to learn, and less likely to drop out of school (Alexander, Entwisle, & Thompson, 1987; Gándara, 1995; Quint, 2006; Stanton-Salazar & Dornbusch, 1995). Positive or supportive relationships between students and staff have been found to be associated with positive outcomes like improved student achievement (Akey, 2006; Martin & Dowson, 2009), while social support married with higher academic expectations has been associated with substantial increases in engagement and achievement (Lee & Smith, 1999; Sebring et al., 1996). Additionally, positive relationships and caring reported by students have been associated with the teachers' value-added scores (MET Project, 2010, 2012).

We focus on these five categories of policies and practices in our analysis of each high school reform model's effectiveness for three main reasons. First, they are part of each model's design. Second, the extant research demonstrates their positive association with our outcome measures – enrollment and/or success in college preparatory mathematics courses. Finally, we focus on these school-wide policies and practices as they should have a direct impact on students' school experiences, rather than on other policies and practices (e.g., those related to leadership or teacher professional development) that may have indirect impacts on students.

To summarize the research above, each of these five policies has been shown to have positive associations with either increased student achievement, motivation, and engagement or improved academic behavior. We hypothesized that the combined effects of all these five policies implemented together may be stronger than the effects of each of these individual policies. Some of the research also leads us to hypothesize that there may be certain interactions between effects of these policies. It has been found that simply increasing course-taking requirements may increase the number of students who take these courses but not necessarily improve their mathematics achievement. At the same time, high academic expectations accompanied by strong academic support and personalized environment and positive relationships with both adults and peers have been shown to improve student outcomes. We hypothesized, therefore, that course-taking, academic support, and personalization may interact to produce a stronger impact on students than course-taking alone.

According to the model design and our theory of change as represented in Figure 1, to accomplish the goal of increasing the number of students who graduate from high school ready for college and work, schools are expected to implement these categories of policies and practices simultaneously. Insofar as these categories of policies and practices are implemented in the context of a whole school reform model, they can be seen as mediators between adopting a reform model and achieving student outcomes. The high level of implementation of these measures is thus seen as a possible mechanism by which the school reform model accomplishes its goals. We recognize, of course, that there are likely other factors unaccounted for in our theory of change. One of the most prominent of these factors is the possibility of students self-selecting into both the ECHS and RHS models. The implications of this are considered in the discussion section.

Thus, according to this theory of change, two conditions lead to the positive impacts of the reform models. First, the policies and practices included in the reform design need to be proven to be effective in improving student learning. Second, these policies and practices should be implemented with high fidelity by schools adopting the reform. Based on the results of the previous studies, we anticipated that we would find positive associations between high levels of implementation of the policies and practices with the successful student pipeline progression. Specifically, we predicted that the combination of all five policies will have a stronger effect than that of individual policies, and that there will be additional effects of interaction among specified policies. Based on the design of the models, on previous research on reform implementation in general and on implementation of these reform models in particular, we predicted that we will see variations in the level of implementation of these policies and practices among selected reform models. This paper evaluates both conditions necessary for the positive impacts of the reform models. The specific research questions are:

1. How do the three high school reform models differ from each other and from traditional high schools in the implementation of the five categories of policies and practices?
2. How are the five categories of policies and practices associated with students' progression in the college preparatory mathematics pipeline through the 10th grade?

Data and methodology

The larger study sought to assess each reform model's success at increasing the percent of students progressing satisfactorily through a pipeline (or sequence) of college-preparatory courses. Viewing a sequence of courses as a pipeline emphasizes the importance of *when* a student takes a course and not just *if* she takes the course. The overall effectiveness of the models was examined elsewhere (Miller & Corritore, 2011; Miller & Mittleman, 2011, 2012). Here, we seek to delve deeper into the models themselves to better understand the mechanisms through which they impact student pipeline progression. Student-level performance data are supplemented by data on the policies and practices collected through surveys administered to principals at reforming and comparison schools.

We developed a survey to measure the extent to which the specific strategies as shown in Figure 1 for each set of policies and practices are implemented by schools. The survey was administered to school administrators (the majority of whom were principals) at all the reforming schools plus a sample of carefully matched non-reforming schools in the summer and fall of 2009.[1] Given that they are usually responsible for policies implementation, principals ought to have the best knowledge of school-wide policies and practices in place at their school. Eighty-three schools responded to our survey (73% response rate). The demographics of the responding schools are described in Table 1.

Table 1. Demographic characteristics of students in schools completing the survey.

Characteristic	ECHS (N = 33)	Redesign (N = 24)	HSTW (N = 10)	Control (N = 16)
Percent minority (%)	38^{a}	$63^{d,b}$	29^{a}	41
Percent eligible for free/reduced-price lunch (%)	41	50	47	41
Average number of students in the school	$133^{a,b,c}$	$296^{b,c}$	919	1026
Percent female (%)	57^{c}	54	50	49

Notes: [a]Significantly different from Redesign schools, $p < .001$; [b]Significantly different from HSTW schools, $p < .01$; [c]Significantly different from control schools, $p < .001$; [d]Significantly different from control schools, $p < .01$.

The survey captures five categories of policies and practices theorized to influence student on-track pipeline progression: course-taking requirements, rigorous instruction, relevance, academic support, and personalization. Each of these broad categories of policies and practices consisted of more specific strategies (indicators) which were identified by theory and prior research. Principal component analyses identified correlations among responses which confirmed that our conceptually based indicators form statistically coherent scales. Table 2 provides a detailed description of the five major categories of policies and practices and their specific strategies. We calculated a value for each category equal to the weighted average of the strategies (most of which are scales themselves).[2] Cronbach's alpha for each policy category and specific strategy is also provided in Table 2. Linking these responses to the student-level data on on-track mathematics pipeline progression yielded a student sample of about 9000 students attending those schools.

Table 2. Specific strategies for the five main categories of policies and practices and their descriptions.

Policy Category	Specific Strategy	Variable or Scale Description
Course-taking Requirements (α = .73)	Math Graduation Requirements	The number of courses in the College Prep sequence (Algebra 1, Geometry, Algebra 2, a course beyond Algebra 2) required for all students before they graduate.
	Ninth-Grade Math Requirements	Whether below-grade students are required to take at least Algebra 1 in the ninth grade.
	Science Graduation Requirements	The number of science courses in the College Prep sequence required for all students before they graduate.
Relevance (α = .76)	Career Relevance (α = .70)	The number of students participating in activities such as concentration of courses in a specific area, career academies or pathways, vocational courses, work study programs, etc.
	College Advisory (α = .86)	The number of students participating in activities such as advising on courses and skills needed for college, college expectations, admissions process.
Personalization (α = .85)	Staff–Student Relationships (α = .79)	Students are well known by staff; there is a mutual respect between staff and students; staff meets with and advises students.
	Staff–Parent Interaction (α = .62)	The frequency and ease of communication between staff and parents.
	Peer and Social Support (α = .70)	Availability of social and emotional support programs and promotion of peer connections and mediation.
Academic Support (α = .75)	Math/Science Specific Support (α = .70)	Whether math- and science-specific supports (tutoring, additional classes) are available and/or mandated for students who need them.
	General Academic Support (α = .65)	Whether general supports (tutoring, additional classes, extended time) are available and/or mandated for students who need them.
Rigorous Instruction (α = .88)	Project-Based/Inquiry Learning (α = .75)	Whether the policy was adopted school-wide or department-wide, or was a decision of each individual teacher. Included practices such as project-based, problem-based, inquiry learning and frequent use of fomative assessment.
	Higher Order Thinking/ Communication (α = .86)	Whether the policy was adopted school-wide or department-wide, or was a decision of each individual teacher. Included practices such as focus on higher order thinking, reasoning, proof, questioning, communication, collaborative learning, and basic skills.

Student-level longitudinal data allow us to observe course-taking and performance as students progress through high school. These data are collected by the North Carolina Department of Public Instruction and housed at the North Carolina Education Research Data Center. These data provide information on student demographic characteristics (e.g., gender, race/ethnicity, and eligibility for free or reduced-price lunch) and prior academic preparation (i.e., eighth-grade test scores, taking Algebra 1 prior to high school, and prior grade-level retention). Course performance is measured on the state-mandated end-of-course exams. The sample includes all students who remained in the same high school through the 9th and 10th grade for the 83 schools for which we have survey data.

We define the mathematics pipeline as consisting of three courses: Algebra 1, Geometry, and Algebra 2 in Grades 9 to 11 to correspond with the North Carolina Department of Public

Instruction’s (NCDPI) graduation requirements for the College Preparatory (CP) Course of Study and admissions guidance for the University of North Carolina system. Successful pipeline progression includes two components. The first component is course-taking and requires that the student take an additional pipeline course each year. The second component is mastery of the subject matter, defined as scoring at the proficient level on the statewide end-of-course exam. Therefore, to be on track a student must have taken and demonstrated mastery in at least one of the three mathematics pipeline courses by the end of 9th grade, and in at least two courses by the end of 10th grade. The survey data reflect the policies and practices in place during the 2008–09 school year. At that time, the last cohort of 9th graders we follow in the larger study were in the 10th grade. Hence, we assessed the relationship of the policies and practices with mathematics pipeline progression at the 10th grade, as we had the most complete data for these students.[3]

To answer our first research question concerning the level of implementation among the three models, we compared the average rating for each of the five scales as well as the individual strategies across the four types of schools (the three reform models plus comparison). Implementation differences were detected using multivariate analyses of variance (MANOVA) with follow-up independent sample *t* tests, adjusted for multiple comparisons using the Benjamini-Hochberg (BH) method (Benjamini & Hochberg, 2000). The effect sizes for differences between means (Cohen’s *d*) were calculated with a standard procedure of difference between means divided by a pooled standard deviation for the two samples.

To explore the relationship between the school-level implementation of the policies and practices and students’ mathematics pipeline progression, we estimate a series of two-level logistic regression models with students nested in schools (Raudenbush & Bryk, 2002). As shown in Figure 2, the probability that student *i* in school *j* is on track at the end of the 10th grade is a function of student characteristics, X^c_i , at Level 1 and the school size, school type (indicator variables for ECHS, RHS, and HSTW), and implementation of the policies and practices at Level 2. A school’s size is equal to the total number of students enrolled.

Figure 2. Two-level logistic model of on-track pipeline progression.

Level 1 Model (student)

$$Pr(Y_{ij} = 1) = P_{ij}$$

$$ln[P_{ij}/(1 - P_{ij})] = \beta_0 + \sum_c \beta_c X_i^c$$

Level 2 Model (school)

$$\beta_0 = \gamma_{00} + \gamma_{01} Size_j [+\gamma_{02} ECHS_j + \gamma_{03} RHS_j + \gamma_{04} HSTW_j] + f(Policy_j) + u_{0j}$$

Student characteristics included the students’ demographic and socioeconomic background as well as proxies for high school academic achievement measured prior to enrolling in high school. All the control variables have been shown elsewhere in the literature to be correlated with academic course-taking and test performance and are related to schools’ decisions to undertake reform initiatives such as those examined here. Their inclusion helps account for the non-random assignment of students to schools to the extent assignment occurs on these observable characteristics. Student demographic and socioeconomic background variables are race/ethnicity (White, Black, Hispanic, and other), gender, eligibility for free or reduced-price lunch, parental education (did not graduate from high school, high school graduate, completed some college, and earned at least a Bachelor's Degree). Measures of prior academic achievement include the students' scores on North Carolina's statewide eighth-grade mathematics exam, whether the student took Algebra 1 prior to high school, and whether the student was over-age when enrolled in the ninth grade (an indicator of probable prior grade-level retention). We also include an indicator of whether the student is observed in the eighth grade in North Carolina (a proxy of mobility immediately prior to high school).

To evaluate our predictions about separate and joint effects of policies and practices, we estimate two models that differ in the functional form of the relationship between policies and practices and student progression. The first includes only the main effects for each policy category individually and combined, and assumes each works independently of the other. The second tests the combined effects of all five policies and the three-way interaction effect between math course requirements, academic support, and personalization (plus the implied two-way interactions), to test our hypotheses specified in the Theoretical background section. These models are estimated with and without the school type indicator variables, which allows us to examine whether the strength of the association of the policy with on-track progression is at least partly a function of its being part of a specific school reform model.

Using the estimated coefficients from the logistic models, we convert the predicted log odds ratios to predicted probabilities of a student being on track associated with a school's levels of policies implementation. We present results for an average student, that is, a student assigned the analytic student sample mean value of all the student characteristics included in the model (e.g., demographics, average eighth-grade test score, etc.). For each policy or set of policies, we compare the average student's predicted on-track probability in two hypothetical schools: one school where the policies are implemented to a high degree and another school where the policies are implemented to a lesser degree. We define a high degree of implementation to be the value at the 75th percentile on the distribution of the policy category ratings among the 83 schools in our sample. A low degree of implementation is the value at the 25th percentile. In each comparison, we alter the degree of implementation of the specific policies in which we are interested while keeping each school at the same size (the average across the 83 schools) and all other policies at the 25th percentile. Standard errors around these on-track probability differences are calculated by applying the delta method to the difference using the estimated variance-covariance matrices (Ai & Norton, 2003). Because the maximum difference in percentage points is determined by the predicted probability in the low implementation school, we also present the difference as the percentage change in the risk of being off track, to facilitate comparisons across the various policy combinations.

Finally, the nature of the pipeline measure of student achievement requires that we observe the same student over time. For a student to be on track in the 10th grade, we need to observe their course-taking and test performance in both the 9th and 10th grades. To account for non-random sample attrition (e.g., due to changing schools) which may bias the results, we assign each student a propensity score weight (Brunell & DiNardo, 2004; DiNardo, 2002; DiNardo, Fortin, & Lemieux, 1996; Robins, Hernan, & Brumback, 2000). Each weight is equal to the inverse of the student's propensity to be censored (i.e., lost to follow-up) multiplied by the inverse of the student's propensity to separate from his initial 9th-grade school.[4]

Results

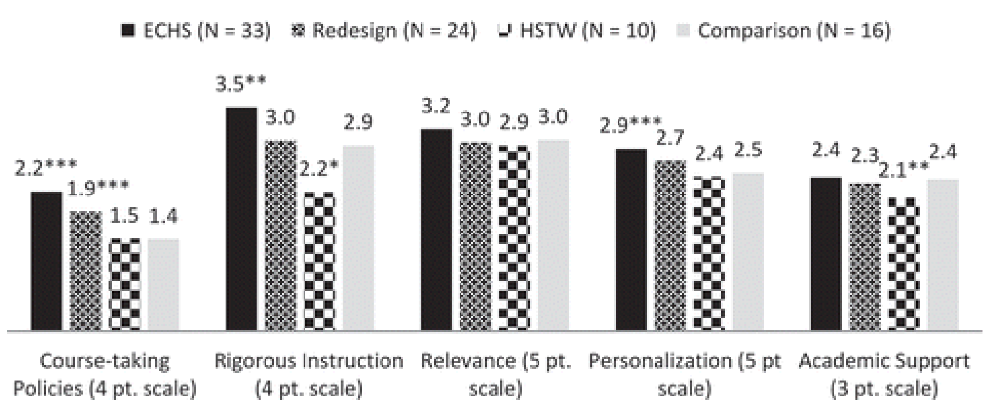
The strength of the survey data is that they begin to help us to unpack the mechanisms through which the reform models influence student pipeline progression. Results we have reported elsewhere indicate that ECHS had a significant positive impact on students' course-taking and performance through the mathematics pipeline and narrowed the group differences in on-track progression rates across race/ethnicity, parental education, and eighth-grade mathematics test scores (Miller & Corritore, 2011). RHS had a positive impact on students' course-taking in mathematics but no impact on their successful performance in those classes. There is also some evidence to suggest that RHS may have reduced gaps in pipeline progression between students who were initially low performing and those who were not (Miller & Mittleman, 2011). Finally, HSTW had no significant impact on students' progression through the mathematics pipeline and may have widened gaps for some subgroups (Miller & Mittleman, 2012).

Some of these differences in impact may be attributable to differences in implementation of the specific policies and practices, or they may be attributable to other, unmeasured differences between the models. The analyses reported in this paper are designed to explore the extent to which these policies and practices differ among the models and the extent to which any differences in implementation are associated with changes in student outcomes. Results from our exploration of the implementation and impact of the policies and practices among the three reform models are organized by the two research questions.

Implementation variation across reform models

The analyses reveal significant differences between types of schools in the implementation of four of the five categories of policies, with only one category, relevance, showing no significant differences. Figure 3 shows the means for implementation of the five policy areas across the different types of schools.

Figure 3. Mean scale scores for policy categories by model.



Implementation of these five categories of policies varies significantly by model, with ECHS implementing three of the policy categories at a higher level. Among the reform models, ECHS have the highest level of implementation of the selected policies, and HSTW have the lowest level of implementation. Both ECHS and RHS were significantly higher ($p < .001$) than control schools on Course-taking Requirements. ECHS are also significantly higher than control schools on Personalization ($p < .001$) and Rigorous Instruction ($p < .01$). At the same time, HSTW were significantly lower than control schools on Rigorous Instruction ($p < .05$) and on Academic Support ($p < .01$). Despite the fact that we collected information on policies from principals only, the degree of variation among the means for policies and practices between different schools suggests that principals relatively accurately presented the extent of their implementation in the schools.

In order to determine whether the broader categories of policies adequately reflect the differences among schools or if we need to consider more specific strategies, we conducted a series of tests to determine if any of the more specific policies and practices composing the broader categories are implemented at varying levels in schools. We examined the difference in implementation between the reform models and the control schools and between the ECHS and all other schools. Table 3shows the means for implementation of all of the specific strategies, and Table 4 shows the effect sizes for the differences between these means.

Table 3. Means of the specific strategies in the five policy categories.

Policy Category	Specific Strategy	ECHS ($N = 33$)	Redesign ($N = 24$)	HSTW ($N = 10$)	Control ($N = 16$)
Course-taking Requirements	Math Graduation Requirements	3.64^{+++}	2.54^{***}	1.40^{***}	2.00^{***}
	Ninth-Grade Math Requirements	0.97^{++}	0.58^{**}	0.60	0.56^{**}
	Science Graduation Requirements	1.94	1.74	1.70	1.63
Relevance	Career Relevance	2.78	2.78	2.67	2.71
	College Advisory	4.21^{++}	3.51^{**}	3.28^{**}	3.53^{**}
Personalization	Staff–Student Relationships	3.28^{+++}	$3.01^{**,+++}$	2.66^{***}	2.60^{***}
	Staff–Parent Interaction	2.44	2.31	2.33	2.49
	Peer and Social Support	2.39	2.22	1.98	2.19
Academic Support	Math/Science Specific Support	2.44	2.42	$2.12^{*,+}$	2.52
	General Academic Support	2.25	2.09	1.93^{***}	2.13
Rigorous Instruction	Project-Based/Inquiry Learning	3.30^{+}	2.58^{**}	$1.98^{+,***}$	2.69^{*}
	Higher Order Thinking/ Communication	3.67^{+}	3.31	2.34^{***}	3.08^{*}

Note: * ($p < .05$), ** ($p < .01$), and *** ($p < .001$) specify significance of differences between ECHS and all other types of schools; + ($p < .05$), ++ ($p < .01$), and +++ ($p < .001$) specify significance of differences between Control and all other types of schools.

Table 4. Effects sizes of the differences between the means of component policies.

Policy Category	Specific Strategy	ECHS vs. Redesign	ECHS vs. HSTW	ECHS vs. Control	HSTW vs. Control	Redesign vs. Control
Course-taking Requirements	Math Graduation Requirements	.98	2.36	1.47	ns	ns
	Ninth-Grade Math Requirements	1.04	ns	1.08	ns	ns
	Science Graduation Requirements	ns	ns	ns	ns	ns
Relevance	Career Relevance	ns	ns	ns	ns	ns
	College Advisory	.85	.91	.93	ns	ns
Personalization	Staff–Student Relationships	.84	1.41	2.00	ns	1.28
	Staff–Parent Interaction	ns	ns	ns	ns	ns
	Peer and Social Support	ns	ns	ns	ns	ns
Academic Support	Math/Science Specific Support	ns	.76	ns	−1.06	ns
	General Academic Support	ns	1.19	ns	ns	ns
Rigorous Instruction	Project-Based/Inquiry Learning	0.75	1.65	0.78	−0.99	ns
	Higher Order Thinking/ Communication	ns	1.63	0.82	ns	ns

An examination of Table 3 shows that the trends identified in the previous analyses continue, with ECHS exhibiting significantly higher levels of implementation on eight of the specific strategies and HSTW exhibiting significantly lower implementation on two strategies. Table 4 reveals large effect sizes for many of the differences between means, ranging from .75 to 2.36. When we look within the broader policy categories, in three instances specific strategies appear to be driving the differences between school types (course-taking, personalization, and relevance); in the other two instances, all of the individual strategies reflect the same trend as the overall category (academic support and rigorous instruction).

Again, one of the goals of this analysis is to test whether variation in school-level implementation is associated with variation in student pipeline progression. In response to the lack of variation in some strategies, we adjusted our measure of three of the five broad policy categories. We constrain our measure of course-taking policies to Math Graduation Requirements and requirements for incoming ninth graders who are below grade in math. Staff–Student Relationships drives the observed differences in the Personalization scale, and we narrow our focus to this strategy. Finally, the broader Relevance scale is replaced by the school's College Advisory strategies.

In the previous section, we determined that the strength of implementation of reform policies varies with reform model type. Specifically, the ECHS model has overall higher implementation and the HSTW model has lower implementation on a variety of policies. In the next section, we examine the associations between implementation of various combinations of policies and students' performance in the college preparatory mathematics pipeline.

Relationship with student pipeline progression

As we noted above, results reported elsewhere demonstrate the effects of the three reform models on a student's probability of being on track in the mathematics pipeline (Miller &

Corritore, 2011; Miller & Mittleman, 2011, 2012). The results reported in the previous section show that specific policies and practices vary across the models; this suggests that the implementation of these policies and practices might be one of the mechanisms by which some models are having a positive impact on pipeline progression. Our second research question is designed to explore the mediating role of specified policies and practices in improving student performance in college preparatory mathematics courses.

We first looked at the extent to which each of the five examined policies/practices is individually associated with students' performance. We then looked at the associations of combinations of policies with students' outcomes. Because all five policy categories are implemented together to some extent in each reform model, we anticipated certain correlations among them. These correlations are reported in Appendix 1. As expected, most correlations between the policies were significant, with some moderate and some low correlations. These correlations justified analyzing the effects of combinations of policies.

In Table 5, we present results from the three statistical models that show the difference in the probability of being on track when the implementation of individual or all five policy areas are at high compared to low levels. The results are organized by the logistical regression model specification. The first column details which of the policies are switched from the low (25th percentile) to the high (75th percentile) ends of implementation distribution. The second column shows the average student's predicted on-track probability in a low implementation school, while the third column shows the student's on-track probability in a high implementation school (with respect to the specific set of policies listed in the first column). The fourth column shows the difference in percentage points between the two previous columns, which is our measure of the association between the policies and pipeline progression. The last column presents this difference as a proportion of the gap between the baseline probability (the second column) and 100%.

Table 5. Difference in predicted probability of being on track in a high- and low-implementation school by policies categories and logistic regression functional form.

Policies increased from the 25th to the 75th percentile of the ratings distribution	On-Track Probability with all at 25th percentile (%)	On-Track Probability at 75th percentile (%)	Difference (percentage points)	Percent of the Gap Closed (%)
Policies' Functional Form: Main effects only				
Rigorous Instruction only	51.1	53.0	1.9 (4.59)	3.9
Academic Support only	51.1	50.6	−0.5 (5.31)	−1.0
College Advisory only	51.1	58.6	7.5 (5.47)	15.4
Math Graduation Requirements only	51.1	55.2	4.0 (7.22)	8.3
Relationships only	51.1	53.5	2.4 (5.23)	4.9
Five policies combined, main effect (no interactions)	51.1	66.0	14.9 (8.17)[+]	30.4
All five policies, Main effects plus controls for the type of reform (no interactions)	56.0	60.0	4.0 (8.93)	9.0
Policies' Functional Form: Main effects plus three-way interaction between math course requirements, personalization, and academic support (plus the 3 two-way interactions)				
Model A1: All five policies	52.1	69.3	17.2 (7.94)*	35.9
Model A2: Model A1 plus controls for the type of reform	56.9	62.7	5.8 (8.36)	13.5

Note: [+]($p < .10$); *$p < .05$.

For example, in Model A1, the average student is predicted to have a 52.1% probability of being on track (and a 47.9% gap with an ideal of 100% being on track) in a school with a low degree of implementation of all five policy categories. However, if this same student attends a similar school with a high degree of implementation of all these policies, the student has a 69.3% predicted on-track probability. The 17.2-point difference in probabilities of being on track between high- and low-implementation schools represents a 35.9% reduction in initial 47.9% gap (17.2/47.9 = .359).

The results in the first panel of Table 5 indicate that if the school type is not taken into account, the influence of each individual policy (holding all of the other policies at 25%) is not significant.[5] On the other hand, greater implementation of the five policies combined is significantly associated with the higher probability of students being on track in college preparatory math courses, especially when interactions between policies are included. The relationship between the level of implementation of policies and on-track probability decreases and becomes insignificant when the school type is included in the regression. In regressions with school type included, when we move the student from a low-implementation school to a high-implementation school we assume both schools are of the same reform type (e.g., both ECHS or both control). These results suggest that there is more variation in the reported implementation of these five policies between types of schools than within each type. Nevertheless, even when controlling for reform type, all differences are in the positive direction. Care should be taken here. This does not necessarily mean that it is the reform model and not the individual policies that is correlated with student pipeline persistence. We showed earlier that some types of reform do a better job of implementing all five policies than other types. In our sample, however, we do not have sufficient numbers of high- and low-implementation schools within each reform type. Almost all our ECHS are high implementers. Therefore, we are limited in our ability to differentiate between the effects of the five policies and the effects of the reform type (which in addition to the five policies would include other attributes that we do not measure in this study).

There is a disproportionately high number of ECHS in the two highest quartiles of implementation for each of the policies, and especially so for the Math Course-taking Requirements. In fact, close to 80% of all ECHS place in the highest quartile for this policy, and between 36 and 47% of ECHS are in the highest quartile for four other policies.

At the same time, the reform model overrepresented in the lowest quartile of the distribution of ratings is HSTW, which among all other types of schools has the highest percentage of schools in the lowest quartile (40% to 70% for different policies), especially for the Rigorous Instruction. Therefore, varying the ratings of the policies from the low to the high levels without controlling for the reform type is also likely to change the likelihood of the school being one type of reform model (most likely HSTW) to another (most likely ECHS).

Discussion

There are two main findings of this study.

First, the three high school reform models examined in the study implement target policies to various extents. The ECHS model demonstrates the highest level of implementation of these policies and shows a significant and consistent impact on the probability of an average student being on track in the college preparatory mathematics pipeline (Miller & Corritore,2011). The HSTW reform model shows the lowest level of implementation of these policies. As reported elsewhere, HSTW had no effect and RHS demonstrated mixed effects on the probability of an average student being on track in the college preparatory mathematics pipeline, as compared to students in comparison schools (Miller & Mittleman, 2011, 2012). This finding is consistent with research on comprehensive school reform that established that the strength of implementation of design features varies significantly among reform models (Berends et al., 2002; Kirby et al., 2001; Supovitz &Weinbaum, 2008; Weinbaum & Supovitz, 2010).

Second, when we examine the influence of the individual target policies (holding all of the other policies constant at the low level), there is no significant association with students' on-track performance. On the other hand, when all policies are jointly implemented at a high level, the policies do have a significant relationship to the probability that an average student is on track in the college preparatory mathematics pipeline at the end of the 10th grade. Compared to the schools where the policy implementation is in the lowest quartile of the ratings' distribution, students in the schools with policy implementation in the highest quartile of the distribution have a probability of being on track in the college preparatory mathematics pipeline that is 17.2 percentage points higher. We cannot interpret our findings to mean that implementing each individual policy or practice separately will not result in change. Although we tested the statistical impact of changes in each of the policies and practices individually while controlling for other policies, in reality all studied schools are implementing all of these policies and practices at some level, making it hard to separate effects of individual policies. Our results do suggest, however, that combinations of these policies and practices are more strongly associated with better student outcomes than individual policies. Including the interaction effect between course-taking requirements in math, academic support, and personalization makes this association with the student outcomes even stronger, supporting our hypothesis that increased course-taking requirements complemented by academic support and personalization provide more beneficial conditions for increasing student success in the college preparatory math course pipeline.

Taken together with the other results from this project reported elsewhere, these results establish the three important links depicted in Figure 1: (a) that the three reform models have differential effects on student outcomes, specifically on student successful progression through the college preparatory mathematics pipeline; (b) that the three reform models have different levels of implementation of select policies and practices; and (c) that combination of selected policies is strongly associated with student successful progression through the college preparatory mathematics pipeline. These results are consistent with previous studies on whole school reform, which found that stronger model implementations of effective practices led to stronger student achievement results (Berman & McLaughlin, 1978; Datnow et al., 2000; Datnow et al., 2005).

Despite the fact that the implementation of policies is measured by school principal self-reports, the results of this study are consistent with the results of prior studies of high school reform models. The significant effects of ECHS on both policy implementation and student outcomes are supported by the results of the ongoing experimental study of the impacts of the model on student academic and behavioral outcomes (Edmunds et al., 2010; Edmunds, Bernstein, Unlu, Glennie, Arshavsky, et al., 2011; Edmunds, Bernstein, Unlu, Glennie, Smith, et al., 2011; Edmunds et al., 2012). These studies report that ECHS have a significant impact on student outcomes. In the 9th and 10th grades, more ECHS than control students successfully completed college preparatory math and science courses, and more ECHS than control students enrolled in college preparatory courses in other core subjects. ECHS students have significantly fewer absences and lower suspension rates, higher continuous enrollment in school through the 10th grade, higher aspirations to attend 4-year college, higher levels of engagement, and more challenging work than students in the comparison group (Edmunds et al., 2010; Edmunds, Bernstein, Unlu, Glennie, Smith, et al., 2011; Edmunds et al., 2012). ECHS students also report higher levels of implementation of specific policies than comparison students, including better relationships with staff, more rigorous and relevant instruction, higher academic expectations, and more academic and social support (Edmunds, Bernstein, Unlu, Glennie, Arshavsky, et al., 2011).

While the ECHS model shows very promising positive outcomes for students, it may not be easy to scale the model up due to its small size and affiliation with a community college or university. In North Carolina, there currently are 76 ECHS, with a few more under development. Due to success of the model in the state, some districts are trying to extend the model into the comprehensive schools, adopting the model's design principles and extending students' access to college courses while in high school. The evaluation of these efforts is currently underway.

Interestingly, while ECHS and RHS reform models in North Carolina have the same design principles and are supported by the same organization (North Carolina New Schools), their implementation of policies and their impacts on student outcomes are significantly different. These differences between ECHS and other types of schools may be attributable to differences in the ways the changes are introduced. In North Carolina, HSTW and RHS reform models are often implemented in existing schools, and thus must work to change the existing school culture. In contrast, ECHS are created as new schools, typically on a community college campus, with newly hired administrators and teachers. Prior to creating a new school, a future principal is engaged in writing a proposal for the new school and in discussion and training with the North Carolina New Schools. The process of conceptualizing the reform activities and then hiring new staff for the school may create more beneficial conditions for staff buy-in into the reform model and its design principles, which was shown in previous studies to affect the strength of implementation (Berman & McLaughlin, 1978; Supovitz & Weinbaum, 2008; Weinbaum & Supovitz, 2010). The strong emphasis in this model is on creating a common vision among staff and a sense of collective responsibility for the common goal of preparing *all* their students to be successful in high school and college despite the students' backgrounds. Additionally, by design ECHS may have more autonomy from the districts and flexibility in using time and resources, which may have contributed to being able to better implement reform features.

ECHS students are expected to earn up to 2 years' worth of college credit free of charge as they earn their high school diploma. This, together with the school's placement on a college campus and extensive supports provided by school staff, creates a very strong incentive for students to succeed academically. In addition, students have to apply to attend the ECHS, and therefore display motivation to work towards a college degree or credit. While many of the ECHS randomly select their students via a lottery, other schools admit students using certain selection criteria. This student self-selection and the school selection process may affect personal characteristics of students admitted to ECHS.

Because HSTW and RHS operate within existing schools with existing staff, they may not experience a uniform buy-in into the changes required by these reform models. These conditions may hinder the fidelity of implementation of the design principles and corresponding policies that comprise these reforms. As a part of this study, we conducted six site visits to get information that would help to interpret our quantitative results. The site visits to two of the HSTW in the study revealed that the schools no longer viewed themselves as implementing the HSTW reform features and were either doing some other type of reform or were not implementing any specific changes. These observations lend support to our survey finding that fidelity of implementation of HSTW design principles may be lower than that in ECHS. It is also consistent with observations from other research that the attention in schools and districts often shifts after a few years from one reform to another (Shiffman, Riggan, Massell, Goldwasser, & Anderson, 2008). The number of visits was unfortunately small due to the budget constraints, so we could not explore this hypothesis further.

In summary, when the three models are compared to each other, there are both similarities and differences between them. The similarities include the measurable background characteristics of the student population served: All three models are targeting underserved populations that include large percentages of minority, low-income, and first-generation college-going students. All three models also support similar policies to be implemented in schools that adopt the model. Additionally, the control schools include students with measurable background characteristics comparable to the reform schools. These similarities allow us to compare these models to one another and to control schools.

At the same time, there are certain differences in the models that allowed us to make predictions about the differences in policies implementation and in student outcomes. ECHS and HSTW models seem to have the biggest differences between their level of policy implementation, with Redesign schools falling somewhere in the middle. The big differences between the ECHS and the two other models are in the school size, in the location on a college campus, and in the way students and staff are selected to be in the school. The average HSTW is 7 times larger than an average ECHS, and 3 times larger than the average Redesign school. At the same time, when we observe the differences between HSTW and control schools, they cannot be explained by the differences in size or population, as these characteristics are similar between the two groups. The two other ECHS characteristics, the location on a college campus and the fact that staff, students, and parents are choosing to apply to these schools, may play a positive role in both the stronger policies implementation and in supporting positive student outcomes. While in this study we statistically control for such observable factors as student background characteristics, including prior achievement, and the level of policies implementation, we may still have some effects of unobservable characteristics such as student and teacher motivation and enthusiasm.

The findings of this study also support the previous research that suggests that if people want to increase the rigor of courses taken by students, changing course-taking policies alone is likely insufficient (Allensworth et al., 2009). Instead, schools need to ensure that the increased course expectations are accompanied by the necessary academic supports (Bloom et al., 2010; Mayer, 2008; Stevens et al., 2008; Swanson et al., 1995) and social supports (Lee & Smith, 1999).

As a final note, our study did find that career relevance policies did not differ among reform and control schools, while the college advisory policies did differ significantly. Despite the HSTW model's stated focus on (a) career relevance, (b) collaboration between vocational and core subject teachers, and (c) workplace mentoring, the career relevance policies are not reported to be implemented more strongly in HSTW than in other reform or comparison schools. This finding may be due to the fact that the HSTW model demonstrated the weakest implementation of all policies in general. The weak reported implementation of reform policies could have led to the weaker student outcomes for this model.

This study presents mostly exploratory analyses of relationships between policies and practices employed by the three reform models and student outcomes. There are several limitations to this study. First, principals were the main source of data, and research has suggested that they may have a limited perspective on certain aspects of schooling. Given that the focus of the study was on school-wide policies, however, the principal appeared to be the best source. An additional limitation concerns the extent to which HSTW schools actually were implementing the model. The limited number of site visits we were able to conduct suggests that schools that are technically considered HSTW by the state or the vendor may not, in fact, consider themselves as implementing HSTW and are likely not implementing the components of the model with any degree of fidelity. The weak implementation of the model, as reported by the principals, may also reflect the fact that this reform was not well sustained and some schools had dropped it. Similar findings for the sustainability of comprehensive school reform efforts were reported by Datnow and colleagues (Datnow, Borman, Stringfield, Overman, & Castellano, 2003). This was not designed as a study of fidelity of implementation of the specific reform models; such a study would definitely be useful (see Edmunds et al., 2013; Supovitz & Weinbaum, 2008).

A final limitation is the fact that ECHS and some RHS were schools that students had to choose to attend; as a result, it is possible the students attending these schools differ on hard-to-measure dimensions, such as motivation, from students in either the control or HSTW schools. This would be a limitation if schools with certain types of students are more likely to implement specific policies and practices than they would if they had a different type of student.

Conclusion

The ECHS, RHS, and HSTW reform models employ certain policies and practices to ensure that all students successfully progress through a college preparatory curriculum in order to be ready for college and work. Consistent with prior research, this study provides evidence that simply increasing course-taking requirements independent of other changes does not have a significant impact. Instead, coupling higher course-taking requirements with other policies related to rigorous instruction, positive relationships between staff and students, and academic support is positively associated with higher likelihood for students to be on track for college in math by the end of 10th grade. However, the study also finds that the three reform models have different levels of implementation of selected policies and practices. The results show that among the reform models, ECHS have the highest level of implementation of all examined policies and the highest association between model implementation and student persistence in the math pipeline.

Previous results have shown that the three reform models have differential effects on student successful progression through the college preparatory mathematics pipeline (Miller & Corritore, 2011; Miller & Mittleman, 2011, 2012). The analyses reported in this paper attribute these effects at least partially to the high level of implementation of selected policies and practices.

Acknowledgements

This research is funded by a research grant (NSF Award 0723412) from the National Science Foundation to The Urban Institute. The opinions expressed are those of the authors and do not necessarily represent views of the National Science Foundation. The authors would like to thank the North Carolina Education Research Data Center (NCERDC) at Duke University for providing the data for this study.

Appendix 1. Correlations among the five selected policies

	1. Math Requirements	2. Advisory	3. Relationships	4. Support	5. Rigor
1. Math Requirements	1				
2. College Advisory	.24*	1			
3. Relationships	.48**	.47**	1		
4. Academic Support	.20	.35**	.26*	1	
5. Rigorous Instruction	.29**	.37**	.39**	.45**	1

Note: * ($p < .05$) and ** ($p < .01$) specify significance level of correlations.

Notes

1. Matched schools were identified using a Mahalanobis distance matching process modeled after Rubin and Thomas (2000). Schools were matched within community type (rural, town, suburb, and city) on a set of variables measured prior to implementation of the reform. These school- level aggregate variables included the percent of students progressing on track through the college preparatory pipeline, student body characteristics (race/ethnicity, parental education, free/reduced-price lunch eligibility), and school enrollment.

2. We predetermined the weights assigned to each indicator based on the theory of the importance of each indicator to the broader policy category. The first author can provide more detail on the weighting process.

3. We stop at 10th grade as many RHS and ECHS were relatively new and we do not have data on a post-reform cohort of 9th graders through the 11th grade. Although the full study focuses on both mathematics and science, we focus only on mathematics in this paper.

4. We use the full-state population (not just our analytic sample) to estimate these propensities.

5. The associations remain insignificant when school type is taken into account. These results are available upon request.

Reference

1. Achieve, Inc. (2004). Ready or not? Creating a high school diploma that counts. Washington, DC: Author.

2. Adelman, C. (1999). Answers in the toolbox: Academic intensity, attendance patterns, and bachelor’s degree attainment. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

3. Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. Washington, DC: U.S. Department of Education.

4. Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. Economics Letters, 80, 123–129.

5. Akey, T. M. (2006). School context, student attitudes and behavior, and academic achievement: An exploratory analysis. New York, NY: MDRC.

6. Alexander, K. L., Entwisle, D., & Thompson, M. S. (1987). School performance, status relations, and the structure of sentiment: Bringing the teacher back in. American Sociological Review, 52, 665–682.

7. Allensworth, E., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College preparatory curriculum for all: Academic consequences of requiring Algebra and English I for ninth graders in Chicago. Educational Evaluation and Policy Analysis, 31, 367–391.

8. Balfanz, R., McPartland, J., & Shaw, A. (2002). Re-conceptualizing extra help for high school students in a high standards era. The Journal for Vocational Special Needs Education, 25(1), 24–41.

9. Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. Journal of Educational and Behavioral Statistics, 25, 60–83. [CrossRef], [Web of Science ®]OpenURL University of North Carolina at Greensboro

10. Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). Facing the challenges of whole-school reform. New American Schools after a Decade (RAND Corporation MR-1498). Santa Monica, CA: RAND Corporation.

11. Berman, P., & McLaughlin, M. W. (1978). Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations (RAND Corporation R-1589). Santa Monica, CA: RAND Corporation.

12. Bloom, H. S., Thompson, S. L., & Unterman, R. (2010). Transforming the high school experience. New York, NY: MDRC.

13. Brunell, T. L., & DiNardo, J. (2004). A propensity score reweighting approach to estimating the partisan effects of full turnout in American presidential elections. Political Analysis, 12, 28–45.

14. Datnow, A., Borman, G., & Stringfield, S. (2000). School reform through a highly specified curriculum: Implementation and effects of the Core Knowledge Sequence. The Elementary School Journal, 101, 167–191.

15. Datnow, A., Borman, G. D., Stringfield, S., Overman, L. T., & Castellano, M. (2003). Comprehensive school reform in culturally and linguistically diverse contexts: Implementation and outcomes from a four-year study. Educational Evaluation and Policy Analysis, 25, 143–170.

16. Datnow, A., Lasky, S. G., Stringfield, S. C., & Teddlie, C. (2005). Systemic integration for educational reform in racially and linguistically diverse contexts: A summary of the evidence. Journal of Education for Students Placed at Risk, 10, 441–453.

17. DiNardo, J. (2002). Propensity score reweighting and changes in the wage distribution (Working paper). Ann Arbor: University of Michigan.

18. DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. Econometrica, 64, 1001–1044.

19. Edmunds, J. A., Bernstein, L., Glennie, E., Willse, J., Arshavsky, N., Unlu, F., … Dallas, A. (2010). Preparing students for college: The implementation and impact of the Early College High School model. Peabody Journal of Education, 85, 348–364.

20. Edmunds, J. A., Bernstein, L., Unlu, F., Glennie, E., Arshavsky, N., & Smith, A. (2011, April). Keeping students in school: Impact of a high school reform model on students' enrollment and progression in school. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

21. Edmunds, J. A., Bernstein, L., Unlu, F., Glennie, E., Smith, A., & Arshavsky, N. (2011, March). The impact of the Early College High School Model on core 9th and 10th grade outcomes. Paper presented at the Annual Meeting of the Society for Research on Educational Effectiveness, Washington, DC.

22. Edmunds, J. A., Bernstein, L., Unlu, F., Glennie, E., Willse, J., Smith, A., & Arshavsky, N. (2012). Expanding the start of the college pipeline: Ninth-grade findings from an experimental study of the impact of the early college high school model. Journal of Research on Educational Effectiveness, 5, 136–159.

23. Edmunds, J. A., & McColskey, W. (2007). Levers for change: Southeast Region state initiatives to improve high schools (Issues & Answers Report, REL 2007–No. 024). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

24. Edmunds, J. A., Willse, J., Arshavsky, N., & Dallas, A. (2013). Mandated engagement: The impact of early college high schools. Teachers College Record, 115(7).

25. Finn, J. D., Gerber, S. B., & Wang, M. C. (2002). Course offerings, course requirements and course-taking in mathematics. Journal of Curriculum and Supervision, 17, 336–366.

26. Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. Educational Evaluation and Policy Analysis, 22, 241–254.

27. Gándara, P. (1995). Over the Ivy Walls: The educational mobility of low-income Chicanos (SUNY Series, Social Context of Education). Albany: State University of New York Press.

28. Jacob, B. A. (2001) Getting tough? The impact of high school graduation exams. Educational Evaluation and Policy Analysis, 23, 99–121.

29. Karp, M. M., Calcagno, J. C., Hughes, K. L., Jeong, D. W., & Bailey, T. R. (2007). The postsecondary achievement of participants in dual enrollment: An analysis of student outcomes in two states. Minneapolis: University of Minnesota, National Research Center for Career and Technical Education.

30. Kaufman, P., Bradby, D., & Teitelbaum, P. (2000). High Schools That Work and whole school reform: Raising academic achievement of vocational completers through the reform of school practice. Retrieved from http://www.sreb.org/page/1386/high_schools_that_work_and_whole_school_reform.html

31. Kemple, J. J. (2008). Career Academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood. New York, NY: Manpower Demonstration Research Corporation.

32. Kirby, S. N., Berends, M., & Naftel, S. (2001). Implementation in a longitudinal sample of New American Schools: Four years into scale-up (RAND Corporation MR-1413). Santa Monica, CA: RAND Corporation.

33. Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. Review of Educational Research, 76, 275–313.

34. Lee, V. E., & Burkam, D.T. (2003). Dropping out of high school: The role of school organization and structure. American Educational Research Journal, 40, 353–393.

35. Lee, V. E., Croninger, R. G., & Smith, J. B. (1997). Course-taking, equity, and mathematics learning: Testing the constrained curriculum hypothesis in U.S. secondary schools. Educational Evaluation and Policy Analysis, 19, 99–121.

36. Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. American Educational Research Journal, 36, 907–945.

37. Lillard, D. R., & DeCicca, P. P. (2001). Higher standards, more dropouts? Evidence within and across time. Economics of Education Review, 20, 459–473.

38. Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement and achievement: Yields for theory, current issues, and educational practice. Review of Educational Research, 79, 327–365.

39. Mayer, A. P. (2008). Expanding opportunities for high academic achievement: An International Baccalaureate Diploma program in an urban high school. Journal of Advanced Academics, 19, 202–235.

40. MET Project. (2010). Learning about teaching: Initial findings from the Measures of Effective Teaching Project. Retrieved from http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf

41. MET Project. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

42. Miller, L. C., & Corritore, M. (2011). Assessing the impact of North Carolina's Early College High Schools on college preparedness (Working Paper). Charlottesville: University of Virginia.

43. Miller, L. C., & Mittleman, J. (2011). Redesigned for success? The effects of high school Charlottesville: University of Virginia.

44. Miller, L. C., & Mittleman, J. (2012). High Schools That Work and college preparedness: Measuring the model's impact on mathematics and science pipeline progression. Economics of Education Review, 31, 1116–1135.

45. Mitchell, K., Shkolnik, J., Song, M., Uekawa, K., Murphy, R., Garet, M., & Means, B. (2005). Rigor, relevance, and results: The quality of teacher assignments and student work in new and conventional high schools. Washington, DC: American Institutes for Research.

46. National Center for Education Statistics. (2001). Students whose parents did not go to college: Postsecondary access, persistence, and attainment. Washington, DC: U.S. Department of Education.

47. National Center for Education Statistics. (2004). The high school transcript study. Washington, DC: U.S. Department of Education.

48. Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). Authentic intellectual work and standardized tests: Conflict or coexistence? Improving Chicago's schools. Chicago, IL: Consortium on Chicago School Research.

49. Newmann, F. M., Lopez, G., & Bryk, A. S. (1998). The quality of intellectual work in Chicago schools: A baseline report. Chicago, IL: Consortium on Chicago School Research.

50. North Carolina New Schools. (n.d.). Design principles for innovative high schools. Retrieved from http://newschoolsproject.org/our-strategy/design-principles

51. Quint, J. (2006). Meeting five critical challenges of high school reform. Lessons from research on three reform models. New York, NY: MDRC.

52. Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage.

53. Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. Review of Educational Research, 79, 3–38.

54. Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11, 550–560.

55. Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. Educational Psychology Review, 17, 327–362.

56. Rubin, D. B., & Thomas, N. (2000). Combining propensity score matching with additional adjustment for prognostic covariates. Journal of the American Statistical Association, 95, 573–585.

57. Sebring, P. B., Bryk, A. S., Roderick, M., Camburn, E., Luppescu, S., Thum, Y. M., … Kahne, J. E. (1996). Charting reform in Chicago: The students speak. Chicago, IL: Consortium on Chicago School Research.

58. Secada, W. G., Chavez-Chavez, R., Garcia, G., Munoz, C., Oakes, J., Santiago-Santiago, I., & Slavin, R., (1998). No more excuses: The final report of the Hispanic Dropout Project. Washington, DC: Office of the Under Secretary, United States Department of Education.

59. Shiffman, C. D., Riggan, M., Massell, D., Goldwasser, M., & Anderson, J. (2008). Channeling adaptation: The role of design in enactment patterns. In J. A. Supovitz & E. H. Weinbaum (Eds.), The implementation gap: Understanding reform in high schools (pp. 46–67). New York, NY: Teachers College Press.

60. Shkolnik, J., Song, M., Mitchell, K., Uekawa, K., Knudson, J., & Murphy, R., (2007). Changes in rigor, relevance, and student learning in Redesigned High Schools. Washington, DC: American Institutes for Research.

61. Southern Regional Education Board (SREB). (n.d.). HSTW goals for continous improvement. Retrieved from http://www.sreb.org/page/1140/goalsconditions_for_continuous_improvement.html

62. Spencer, V. G. (2006). Peer tutoring and students with emotional or behavioral disorders: A review of the literature. Behavioral Disorders, 31, 204–222.

63. Stanton-Salazar, R. D., & Dornbusch, S. M. (1995). Social capital and the reproduction of inequality: Information networks among Mexican-origin high school students. Sociology of Education, 68, 116–135.

64. Stasz, C., & Brewer, D. J. (1998). Work-based learning: Student perspectives on quality and links to school. Educational Evaluation and Policy Analysis, 20, 31–46.

65. Stevens, D., Sporte, S., Stoelinga, S. R., & Bolz, A., (2008). Lessons from high performing small high schools in Chicago (Research Brief). Chicago, IL, Consortium on Chicago School Research.

66. Stone III, J. R., Alfeld, C., & Pearson, D. (2008). Rigor “and” relevance: Enhancing high school students’ math skills through career and technical education. American Educational Research Journal, 45, 767–795.

67. Supovitz, J. A., & Weinbaum, E. H. (2008). The implementation gap: Understanding reform in high schools. New York, NY: Teachers College Press.

68. Swanson, M. C., Mehan, H., & Hubbard, L. (1995). The AVID classroom: Academic and social support for low-achieving students. In J. Oakes & K. H. Quartz (Eds.), Creating new educational communities. Ninety-fourth yearbook of the National Society for the Study of Education. Part I (pp. 53–69). Chicago, IL: University of Chicago Press.

69. Takoko, N., & Allensworth, E. (2009). “Double-dose” Algebra as an alternative strategy to remediation: Effects on students' academic outcomes. Journal of Research on Educational Effectiveness, 2, 111–148.

70. Weinbaum, E. H., & Supovitz, J. A. (2010). Planning ahead: Make program implementation more predictable. Phi Delta Kappan, 91, 68–71.