“In theory it’s a good idea”: Understanding implementation of proficiency-based education in Maine

Karen Shakman, Brandon Foster, Noman Khanani,
Jill Marcus, and Josh Cox

STUDENT-CENTERED LEARNING
RESEARCH COLLABORATIVE

Bridging the worlds of research, practice, and policy, JFF's Student-Centered Learning Research Collaborative investigates student-centered approaches to improve outcomes for learners from all backgrounds, particularly those who have been marginalized or underserved by the current system. This bold initiative began in 2016 with a core group of scholars, school leaders, policymakers, practitioners, and funders—each known for their impact and influence—coming together to clarify and catalyze the field. Since that time, the Research Collaborative has supported:

- multiple research teams employing a diverse set of research methods to build the evidence base for student-centeredlearning;
- a variety of field-advancing projects that accelerate innovation and generate investment in student-centered practices;
- a cohort of Students at the Center Distinguished Fellows who show what's possible when applications of student-centered practices are driven by rigorous research and a commitment to equity;
- and a series of public-facing resources designed to scale implementation and ensure all students flourish in our schools.

Education Development Center (EDC) conducted this study as part of the Research Collaborative's initial cycle of research. The team at EDC worked alongside fellow scholars, educators, and policymakers to investigate the impact of specific student-centered practices and then translate their findings for cross-sector audiences. This report represents their work over the past two years as they partnered with ten districts in rural Maine to examine the extent to which students were exposed to and experienced proficiency-based education.

Other Research Collaborative studies in this cycle include:

- *Implementation of Student-Centered Learning Approaches,* American Institutes for Research
- *Learning With Others: A Study Exploring the Relationship Between Collaboration, Personalization, and Equity*, American Institutes for Research
- *Abolishing the phrase "I'm not a math person",* High Tech High Graduate School of Education

For more information about and additional resources derived from this study from American Institutes for Research and the Student-Centered Learning Research Collaborative, visit sclresearchcollab.org.

This study is generously funded by the Nellie Mae Education Foundation.

Table of Contents

Abstract

Student-centered learning encompasses four overlapping and complementary principles (JFF, 2014): competency-based progression, personalization, flexibility in where and when learning takes place, and facilitation of key skills and dispositions such as agency and ownership. To date, few studies have attempted to quantitatively characterize implementation of student-centered learning in order to investigate the relationship between variability in implementation and student outcomes—particularly outcomes among high-need student subgroups (Steele, Lewis, Santibañez, et al., 2014). Education Development Center (EDC) partnered with 10 districts in rural Maine that were in the process of implementing the state's requirement that students graduate with a proficiency-based diploma, to study students' exposure to student-centered, proficiency-based education and the relationship between exposure and student academic performance and engagement. Using Latent Profile Analysis, a statistical technique used to uncover hidden subgroups (i.e., latent profiles) based on the similarity with which a group of individuals responds to a set of survey questions, we found that three distinct proficiency-based education (PBE) exposure profiles existed, in similar proportions across all the participating schools and within every school. Analyses of district level administrative data showed that having an IEP was associated with higher exposure to PBE practices but that other student characteristics, including free and reduced-price lunch status and gender were not associated with more exposure to PBE practices. We also observed a positive relationship between exposure to PBE practices and increased levels of student engagement, and a negative association between exposure to PBE practices and SAT scores. Finally, qualitative analyses revealed that implementation to date has largely focused on identifying graduation standards and implementing new proficiency-based grading practices, with traditional classroom practices still fairly commonplace.

Introduction

Throughout the nation, interest is growing in using student-centered approaches to learning and instruction to increase graduation rates and ensure postsecondary readiness for all students (Wolfe, Steinberg, & Hoffman, 2013), yet there is currently little rigorous research that investigates how the implementation of student-centered learning (JFF, 2014) relates to student outcomes. In particular, few studies have attempted to quantitatively characterize implementation in order to investigate the relationship between variability in implementation and student outcomes—particularly outcomes among high-need student subgroups (Steele, Lewis, Santibañez, et al., 2014).

Student-centered learning encompasses four key principles (JFF, 2014): competency-based progression, personalization, flexibility in where and when learning takes place, and facilitation of key skills and dispositions such as agency and ownership. In this study, rather than addressing one or some of these principles, we sought to understand and quantitatively characterize the overlapping and complementary nature of all four principles in practice.

In addition, although there is an emerging consensus about key principles of student-centered learning, implementation of associated reforms varies substantially across states and school districts (Scheopner Torres, Brett, & Cox, 2015; Steele et al., 2014; Wolfe et al., 2013). For example, previous research on student-centered, competency-based reforms demonstrates that policies can range from reforms that simply allow flexibility in awarding credit to those that completely transform the education system (Sturgis & Patrick, 2010). In addition, although a few studies have attempted to assess the "effect" of student-centered learning, these studies have treated such learning only as a present/not present phenomenon (e.g., Steele et al., 2014). Yet a focus on causal effects of student-centered learning must grapple with the variability in how the key principles of student-centered learning are implemented in practice. In particular, given that these principles are hypothesized as overlapping and complementary, empirical research is needed that characterizes the implementation of practices aligned with each of these principles, including the ways in which the principles do or do not overlap in practice. Addressing these gaps, the Education Development Center (EDC) partnered with 10 districts in rural Maine to study students' exposure to student-centered, proficiency-based education and the relationship between exposure and student academic performance and engagement.

Proficiency-based Education in Maine[1]

As described above, student-centered learning encompasses four core principles, with competency-based progression being among them. At the time of this writing, Maine state law still required that all students graduate from high school with a proficiency-based diploma, however this policy may soon change. Although the terminology in Maine is *proficiency* rather than *competency* as described in the framework above, the goals are the same: to ensure that students have the opportunity to progress as they demonstrate mastery of important content and skills. Our study sought to understand exposure to proficiency-based education (PBE) as a reform effort in keeping with the core principles of student-centered learning.

In 2012, the Maine Legislature passed *An Act to Prepare Maine People for the Future Economy* (S.P.439 – L.D.1422). This law required high school students to show proficiency in eight content areas (Career and Education Development, English Language Arts, Health Education & Physical Education, Mathematics, Science & Technology, Social Studies, Visual & Performing Arts, and World Languages). The law also required Maine students to demonstrate proficiency in the "Guiding Principles", a set of skills that include communication, self-direction, problem solving and thinking skills as well as citizenship. They are a part of the state's K-12 learning standards, established in 1997 and known in Maine as the Maine Learning Results[2]. The 2018 graduating class was originally designated as the first who would receive a proficiency-based diploma in accordance with the law.

Three years later, in 2015, many public school districts were granted extensions, allowing them to move the implementation date back to 2020. A new law, titled *An Act to Implement Certain*

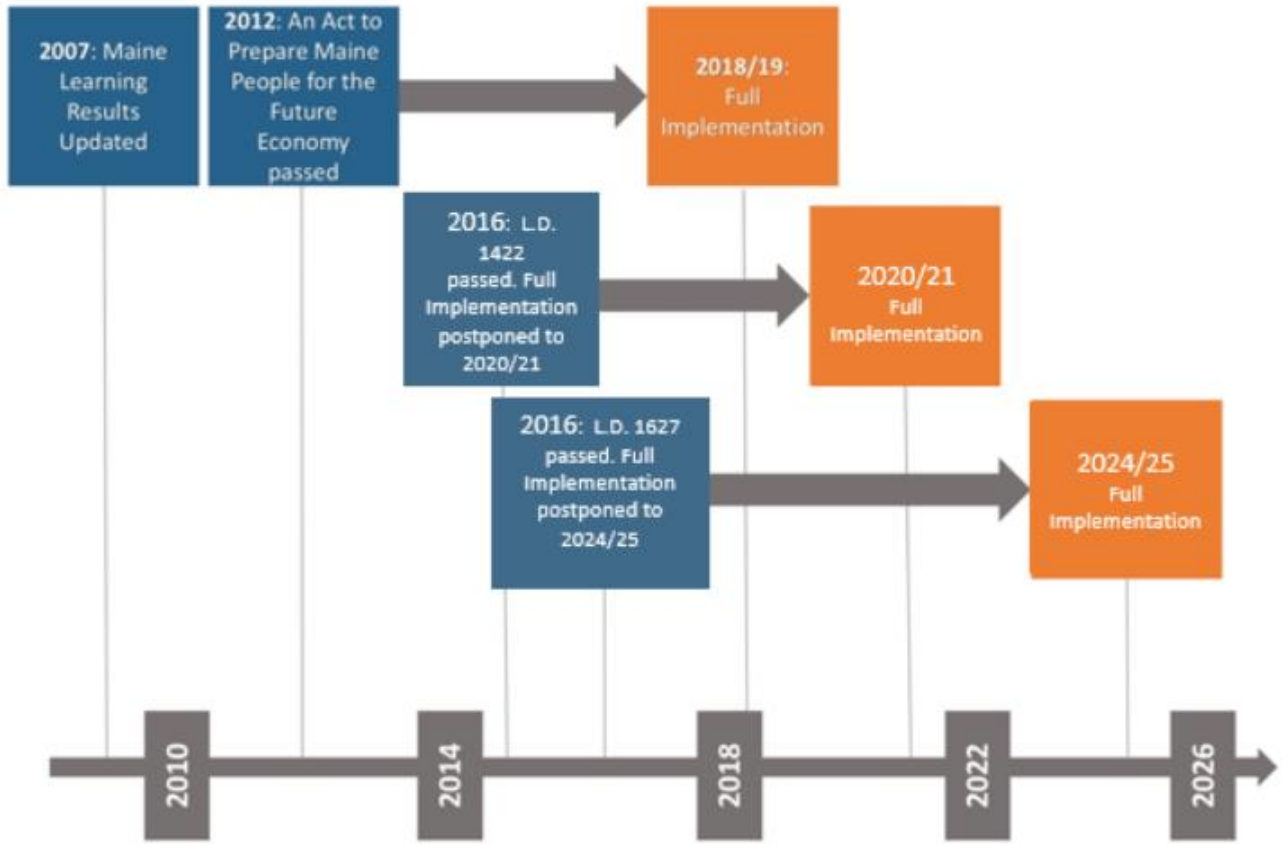
[1] The content for the discussion of Maine's policy context is derived and adapted from the following article: Stump, E., Johnson, A., Jacobs, C. (2017). *Proficiency-based High School Diploma Systems in Maine: Implications for Special Education and Career and Technical Education Programming and Student Populations.* Center for Education Policy, Applied Research and Evaluation in the School of Education and Human Development, University of Southern Maine, Gorham, ME.

[2] According to Stump, et al, 2017: "The *Maine Learning Results: Parameters for Essential Instruction* were reviewed and then updated in 2007, with critical changes to content areas standards and the guiding principles. At that time, legislation was passed requiring the annual state assessments to reflect students' proficiency levels as defined by the updated standards in Mathematics, Reading, and Science. In addition, the updated *Maine Learning Results* were formally integrated within state policies related to school funding and school accountability measures. Although a statewide attempt to require a common local assessment system based on the *Maine Learning Results* standards ended unsuccessfully in this same year, practitioners had dedicated significant time across the past decades discussing standards with students as well as building standards-based curricula and assessments (Leiberman & Miller, 2011; Stump, Silvernail, Fallona & Moran Gunn, 2013; Stump & Silvernail, 2014). In 2011, Maine adopted the Common Core State Standards in Mathematics and English Language Arts. Although state law and the Maine Constitution prohibit a mandatory statewide curriculum, the Maine Department of Education (MDOE) encouraged and supported local efforts to align curricula and assessments to the state-developed *Maine Learning Results*." (pp. 9-10)

Recommendations of the Maine Proficiency Education Council (S.P. 660 – L.D. 1627) enabled schools to implement proficiency-based graduation requirements over a period of time: for the four core content areas by 2020/21 and for all eight content areas and the guiding principles by 2024/25.

Most recently, the state legislature's education committee voted to endorse Bill L.D. 1666, which removes the proficiency-based diploma requirement and revises language in the state law to instead require the issuance of high school diplomas based on students' meeting state standards. Although the proposed changes to the law call for an end to the proficiency-based diploma mandate, school districts are still offered the flexibility to implement proficiency-based education in those communities where it is determined that the approach can best meet the needs of students, parents, and the school community. At the time of the writing of this report, the Maine Legislature has not yet voted on Bill 1666.

Figure 1: Timeline of legislation leading to adoption of proficiency-based high school diploma in Maine



The above timeline outlines key pieces of legislation leading to and following Maine's decision to mandate a proficiency-based high school diploma.

Local Context

EDC conducted the research in collaboration with a consortium of districts in rural Maine. The districts, ranging in size from 150 to 1,800 students, are primarily rural and fairly remote. They do not have access to significant supplemental resources, professional development, or external partners. The training and preparation they have done to date to transition to PBE has been primarily self-managed, either within the district or through attendance at various regional events (the districts are part of a collaborative of districts that participate in common professional development), where external consultants have provided one-day trainings. The consortia of districts originally contacted EDC seeking additional support related to implementation of PBE. Although EDC was not, at the time, in a position to offer professional training or development for educators in how to transition to PBE, the study described in this report has shone a light on areas of strength and places on which the districts may focus their work going forward.

Research Questions

Our study posed the following research questions:

1. To what extent do specific patterns of student exposure to student-centered, proficiency-based education exist?
    a. What are the characteristics of the exposure profiles?
    b. What predicts membership in the exposure profiles?
2. How is student exposure to student-centered, proficiency-based education associated with student academic and engagement outcomes?
    a. Do associations between exposure and outcomes vary as a function of student characteristics?[3]
3. What is the nature of implementation and what factors contribute to variability in the implementation of practices aligned with principles of student-centered, proficiency-based education?

[3] The original research question referenced socioeconomic status and prior achievement as moderators, but the research question was changed to more generally focus on student characteristics. We made this change to reflect the interactions we examined for all student characteristics for which we had data, and not just FRL. We did not have sufficient data for prior achievement to warrant an investigation of the effect of prior achievement on the association between exposure and outcomes.

Study Sample

The sample for the study consisted of students in grades 9–12 who attended 11 secondary schools located within 10 districts in rural Maine during the 2016/17 academic school year. For qualitative data collection, the sample comprised students in grades 9–12 in a subsample of 3 of the 11 schools in the 2017/18 academic school year. This sample was purposively selected to include secondary schools navigating the implementation of PBE and representing varying student body sizes and levels of student economic need.

EDC researchers collected administrative data for 2,270 high school students from the participating districts. Such records included the following elements:

- Student enrollment.
- Grade level.
- Attendance.
- Free and reduced-price lunch (FRL) status.
- Individualized education plan (IEP) status.
- Graduation status for the class of 2017.

More than 51% of the students in the sample were male, 45% received FRL, 18% were on an IEP, and 93% were White. The proportion of students was approximately equal across each grade level (9th–12th), and nearly all members of the class of 2017 (98%) graduated. On average, students missed 6.5 days of school because of unexcused absences during the 2016/17 academic year.

Among this sample of students, 1,828 students responded to the survey used to measure student exposure to PBE and student engagement (80.5%).[4] Table 1 provides student characteristics for both samples compared with national and state averages.

[4] These figures do not include 42 survey respondents for whom administrative data were not available.

Table 1. Student characteristics by sample

Sample	National Average	State Average	Administrative Data Sample	Survey Sample
Number of Students	-	181,599	2,270	1,828
Male	-	51.43%	51.42%	50.33%
FRL	48.10%	45.60%	45.06%	43.10%
IEP	13.00%	16.70%	18.13%	15.89%
Graduated	83.00%	87.50%	97.83%	98.56%
9th Grade	-	25.23%	25.73%	27.90%
10th Grade	-	25.57%	24.45%	24.95%
11th Grade	-	24.71%	25.42%	24.29%
12th Grade	-	24.49%	24.41%	22.87%
White	50.00%	90.84%	92.91%	90.05%
Unexcused Days Absent	-	-	6.48	5.74*
SAT	1060	1012	982.58	988.01*
PSAT	970	957	946.7	945.2
Unweighted GPA	-	-	83.74	84.1

Note. National data source: National Center for Education Statistics, Digest of Education Statistics: 2016. State data source: Maine Department of Education, YEAR. A * indicates statistically significant differences at $p \leq 0.01$ between the survey sample and the administrative data sample. Student characteristics from this survey sample did not statistically significantly differ from the total administrative data sample, with the exception of unexcused days of absence. National average SAT and PSAT data pulled from the College Board website. State average data from the Maine Department of Education website. National and State level GPA data are not available. Most districts recorded GPA on a scale of 0 to 100, with the exception of one district that recorded it on a scale of 1 to 4 and another district that collected only weighted GPAs (0 to 100+ scale). Another district had already established a proficiency-based system, discontinuing the traditional GPA system for their 9th and 10th grades while still maintaining traditional GPA records for 11th and 12th grades. GPA was thus standardized in the regression analyses.

The sample for this study reflects the demographics of the entire state. The only notable difference is the higher graduation rate for the sample compared with the state of Maine and with the nation (nearly 98.6% of the survey sample graduated in 2017, compared with 87.5% of Maine 12th graders in 2015 and 83% nationally). There is also a lower proportion of FRL students (43%) compared with the nation (48%), and a higher proportion of students with IEPs (15.9%) and White students (90%) compared with national averages.

Researchers also collected qualitative data to further understand implementation of PBE. The qualitative component included interviews with district superintendents (in summer 2017), school principals, and teachers; and focus groups with students and their parents or legal guardians (in fall 2017).

EDC selected three high schools to participate in the qualitative component of the study. Throughout this report, the schools are referred to with the following pseudonyms: Princeton, Fields, and Walden. These three schools were selected from among the 11 schools participating in the study based on variation in the following areas:

- Student enrollment.
- Student demographic characteristics.
- Implementation pace/approach.
- Student survey results.

Specifically, we sought to ensure variation in grade 9–12 student enrollment, student demographics including percentage of students who qualify for FRL, status of implementation of PBE,[5] and aggregated student responses to the PBE student survey.

[5] Implementation status was determined based on interviews with superintendents and principals in spring, 2017.

Methodology

Overview of Approach

The current study used a mixed methods design, employing both quantitative and qualitative data collection and analyses. The study aimed to facilitate understanding of implementation of the various and overlapping features of PBE and of the relationship of these features to student engagement and academic outcomes. A description of the quantitative and qualitative methods and measures follows.

Research Question 1: To what extent do specific patterns of student exposure to student-centered, proficiency-based education exist?

EDC, in partnership with the 10 study districts (11 high schools), administered the *Competency-based Learning Survey for Students* (Ryan & Cox, 2016). The *Competency-based Learning Survey for Students* was designed to provide schools and school districts with a resource for measuring student exposure to and understanding of student-centered practices in schools that are implementing a competency- or proficiency-based approach to learning and instruction. The survey, which includes 30 items and collects information on several constructs, is organized into an introductory section and six survey modules.

Module A collects demographic information to help users understand the characteristics of the student population taking the survey. Module B addresses students' beliefs and understanding about competency- or proficiency-based learning. Each of the remaining four modules (modules C-F) collects data on students' exposure to a key element of competency-based learning: progression through demonstration of mastery, personalization, flexible assessment, and development of specific skills and dispositions.

In November 2016, each of the 11 high schools administered the survey to all grade 9-12 students (excluding opt-outs, which ranged from zero to approximately 10 opt-outs per school, and excluding students ineligible to participate because of language or learning differences, which ranged from zero to 24 students). Survey response rates were high in all participating high schools, ranging from 81% to 91% when absent students were treated as nonresponses and from 87% to 99% when absent students were excluded from the calculation.

To answer research question 1a, latent profile analysis (LPA) was used to investigate whether there were patterns in students' exposure to elements of student-centered, proficiency-based education. LPA is a statistical technique used to uncover hidden subgroups (i.e., latent profiles) based on the similarity with which a group of individuals responds to a set of survey questions (McLachlan & Peel, 2000; Vermunt & Magidson, 2004). In this study, LPA was employed to empirically test whether there are subgroups of students, as revealed by their responses on the survey, who share a common experience of exposure to aspects of proficiency-based education (PBE) in their schools. We used student responses to 16 high-leverage survey items to discern the subgroups. We selected these high-leverage items based on a review of the relevant literature coupled with exploratory analyses of the survey results. Further information on the selection of these items is available in Appendix A.

Research Question 1A: What are the characteristics of the exposure profiles?

To understand the results of the LPA and the distinct patterns in student exposure, researchers employed a mixed methods approach. In addition to gathering the survey responses, as described above, EDC conducted qualitative data collection in three districts, selected to represent the full range of schools in the sample in terms of student demographics and implementation status. The qualitative component of the study included the following elements:

- Interviews with district superintendents: In the early summer of 2017, EDC researchers conducted phone interviews with the superintendent at each of the three sites.
- Interviews with school principals: In the early fall of 2017, EDC conducted phone interviews with the high school principal at each of the three sites.
- Interviews with teachers: In November 2017, EDC visited the three high schools and interviewed between three and five teacher participants at each of the schools. These teachers represented different content areas and/or grade levels (see Table B1 in Appendix B for a description of the teacher respondents).
- Focus groups with students: EDC researchers relied on administrators' recommendations to identify six to eight student focus group participants at each of the three schools. These focus groups were conducted on-site at each school during the November 2017 site visit.
- Focus groups with parents/guardians: EDC researchers also relied on administrators' recommendations to identify six to eight parent/guardian focus group participants at each of the three schools. These focus groups were also conducted on-site at each school during the November 2017 site visit.

Researchers used standard qualitative analysis practices to descriptively code and analyze the data from the interviews and focus groups (Glazer & Strauss, 1967; Miles & Huberman, 1994). Researchers generated key themes derived from initial coding of all these data (Charmaz, 2000) to create both

individual case descriptions and cross-case findings (Patton, 2002). In generating the cross-case findings, researchers sought connections between the survey and other quantitative results and the data from the interviews to both confirm and disconfirm results from the quantitative analyses.

Research Question 1B: What predicts membership in the exposure profiles?

After identifying a set of latent exposure profiles for PBE, we used multinomial regression to identify any associations between student characteristics and latent profile membership. The following student characteristics were analyzed:

- IEP status.
- English Learner status.
- Gender.
- Grade level.
- Race (White or non-White).
- Number of unexcused absences in the school year.
- Unweighted GPA.

Research Question 2: How is exposure to student-centered, proficiency-based education associated with student academic and engagement outcomes?

To investigate the association between PBE profiles and outcomes, we used multilevel regression models with students clustered within schools. The goal of conducting these analyses was to understand whether any associations existed between students' exposure to a particular set of PBE practices and those students' academic and engagement outcomes. The two outcomes for the regression analyses were student engagement and academic achievement.

Student engagement survey: We employed a pre-existing, validated survey to measure student engagement. Voelkl's (1996) Identification with School Questionnaire (ISQ) was developed from a theory of student engagement in school called the participation-identification model (Finn, 1989, 1993). The instrument consists of two subscales: "belongingness," defined as the degree to which a student feels that he or she is a significant member of the school community, and "valuing," or the importance a student places on learning and school (Finn, 1989; Voelkl, 1997). The survey is included in Appendix C. As shown in Table 2, the survey exhibited reliability in this sample (Cronbach's alpha = 0.84).

Table 2. Engagement survey item statistics

Scale Statistic	Total Engagement Score
Mean	42.88
Variance	57.657
Std. Deviation	7.593
N of Items	16
Cronbach's Alpha	0.843
Cronbach's Alpha Based on Standardized Items	0.844
N	1785

Note. Students who did not answer all 16 survey items were not included in this analysis.

Academic achievement: The association between student exposure to PBE and student academic outcomes was measured using 11th grade SAT scores. All high school students in the state are required to take the SAT. Scores were available for most students in the survey sample for 11th grade (90%). Average SAT scores within each sample are displayed in Table 1 above.

Research Question 2A: Do associations between exposure and outcomes vary as a function of student characteristics?

To investigate whether the association between PBE exposure profiles and outcomes varied as a function of student characteristics, we built on the multilevel regression models in Research Question 2. We specified interaction terms between each student characteristic variable and each of the PBE exposure profiles. Models were fit separately for students' engagement and SAT scores. Statistically significant interactions in these models can be interpreted to mean that the relationship between PBE exposure profiles and outcomes depends on the characteristics of the students. For example, a positive interaction between free and reduced-price lunch (FRL) and PBE profile membership would support the idea that associations between PBE profiles and outcomes are stronger for students who receive FRL than students who do not. More specific details for the analyses are provided in Appendix D.

Research Question 3: What is the nature of implementation and what factors contribute to variability in the implementation of practices aligned with principles of student-centered, proficiency-based education?

Researchers collected qualitative data from three sites to provide more contextual information about the status of implementation, challenges and successes encountered thus far, and attitudes toward and beliefs about implementation of the PBE reform in each of the three high schools. As described above, the schools were selected to represent the range of schools in the overall sample in terms of demographics, and they reflect the diversity of experiences of implementation to date. Interview questions were designed to complement the quantitative component of the study, particularly related to survey results; questions on the interview protocols mapped to the domains in the survey: attitudes and beliefs about PBE; progression through demonstration of mastery; personalization; flexible assessment; and development of specific skills and dispositions (see Appendix E for the interview and focus group protocols). Interview and focus group protocols were also designed to elicit complementary information from key constituents. In other words, we solicited students' and parents' understanding of what PBE is and why the school has embarked on the reform; and we solicited information from superintendents and principals about how they communicate information about the reform to students and families. These complementary questions permitted us to understand more about the degree to which there was a common understanding among key constituents about the nature and purpose of the reform effort and a shared understanding of the overall goals of the PBE model. Employing traditional qualitative analyses, we reduced the data using thematic analyses and then developed descriptions of each of the three schools that summarize the status of implementation and attitudes and beliefs of key constituents.

Results

Next, results are presented separately for each research question. Information about the latent profiles are found in Research Question 1A and predictors of profile membership are in Research Question 1B. Regression models examining the association between the profiles and the outcomes are provided in Research Question 2. Technical details of the results are provided in Appendix D.

Research Question 1: Specific patterns of exposure. Three unique profiles of students' exposure to PBE practices emerged across the sample and in every school

Based on the results from exploration of 1-, 2-, and 4-profile solutions, we concluded that the appropriate solution contained three PBE exposure profiles: Minimal, Low-Medium, and Medium Exposure. The fit comparisons across the various competing models are summarized in Table A1 in Appendix A, and the absolute measures of model fit are provided in Table A2. Overall, results indicate that not only did the 3-profile solution provide adequate separation of the latent profiles (that is, the 3-profiles solution generated profiles that were distinct from one another), but also students were assigned to their most likely profiles with a high degree of certainty. Summary statistics for the average response to each question in the LPA for each exposure profile are provided in Table 3. Note that the rating scales associated with the survey items are provided in Table A3 of Appendix A.

Table 3. Summary statistics for the PBE exposure profiles

Survey Item	Minimal			Low-Medium			Medium		
	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed
MASTERCOMP: I must show my teachers that I have mastered each standard before I can move on to the next one.	3.0	0.08	24.28%	2.4	0.12	47.12%	1.2	0.09	60.68%
NEXTCOMP: I am able to move on to the next standard when I am ready, even if other students in the course are not ready.	2.6	0.20	10.96%	1.4	0.07	27.96%	0.5	0.05	52.22%

Survey Item	Minimal			Low-Medium			Medium		
	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed
RUBRIC: My teachers give me a rubric so that I know how I am progressing on each standard.	3.1	0.06	35.18%	2.3	0.08	46.94%	1.8	0.10	61.54%
ONLINECOMP: I am able to complete some or all of the course requirements online.	1.2	0.04	15.97%	0.8	0.04	27.17%	0.5	0.05	40.67%
PROJCRED: If I complete a project that wasn't assigned at school but is related to a course I am taking, I can earn credit for the project in that course.	1.1	0.13	3.23%	0.4	0.04	12.27%	0.1	0.03	35.43%
INTERNCRED: I can earn credit for completing an internship or job-shadowing in the community.	1.3	0.10	11.60%	0.8	0.05	26.07%	0.3	0.08	43.97%
GIVELECTUR: My teachers spend most of class time giving a lecture or presentation to the whole class.	2.4	0.09	55.82%	2.5	0.04	49.60%	2.8	0.05	48.44%
SMALLGROUP: My teachers work with students in small groups or individually.	2.4	0.14	19.40%	1.7	0.03	33.88%	1.0	0.05	48.36%
TEACHDIFF: My teachers teach the material in several different ways in order to help students learn.	3.0	0.07	17.34%	1.9	0.07	38.62%	0.9	0.09	55.98%

Survey Item	Minimal			Low-Medium			Medium		
	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed
CHOOSEHOW: I have had opportunities to choose how to show my teachers what I have learned.	2.1	0.10	11.85%	1.2	0.05	29.48%	0.5	0.04	52.85%
RETAKE: If I do poorly on an assignment on the first try, I can try again later.	2.8	0.12	32.02%	2.2	0.11	43.00%	1.6	0.12	55.96%
DEMOLEARN: To show that I have mastered a course standard, I must demonstrate my learning in more than one way.	2.8	0.10	20.00%	1.9	0.06	37.76%	1.0	0.08	55.92%
ADVICE: When I have trouble learning something new, my teachers give me advice and strategies that help me to stick with it.	3.2	0.04	23.74%	2.3	0.08	45.32%	1.2	0.06	64.22%
TRACKPROG: Teachers show students how to keep track of their progress on each of the standards.	3.1	0.06	17.24%	2.1	0.07	42.02%	0.9	0.08	62.32%
ONTIME: Teachers show students strategies for making sure all assignments are completed on time.	3.0	0.05	17.92%	2.1	0.08	42.78%	0.9	0.09	60.88%
PREPGRAD: I know which steps to take during high school in order to prepare for what to do after I graduate.	3.1	0.04	32.44%	2.5	0.07	49.92%	1.6	0.12	62.82%
Sample Size Within Profiles									
N	562			929			332		

Survey Item	Minimal			Low-Medium			Medium		
	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed	M	S.E.	Prop. Endorsed
Proportion of sample	0.31			0.51			0.18		

Note. M = estimated mean of the item for participants assigned to a particular profile. SE = standard errors for the estimated means. Prop.endorsed = the mean response to survey items within each profile divided by the total number of categories in the rating scale (higher values indicate students reporting exposure to more of the practice described in the item).

A profile plot for the three PBE exposure profiles is provided in Figure 2. The figure provides a visual representation of the pattern of students' responses to the 16 high-leverage survey items for students assigned to each of the three profiles. This visual provides a mechanism for observing what is both similar and different in student responses across profiles. In the plot, the y-axis has been transformed to a proportion. This transformation was made to account for the fact that some items used 4-point rating scales and others used 5-point rating scales. This transformation ensured that the comparison across items was fair. The interpretation of the y-axis in the plot is as follows: If the average response to a particular item is further up the y-axis, that indicates that students in that profile report more exposure to the PBE practice associated with the particular item.

Figure 2. Characteristics of the Latent PBE Profiles

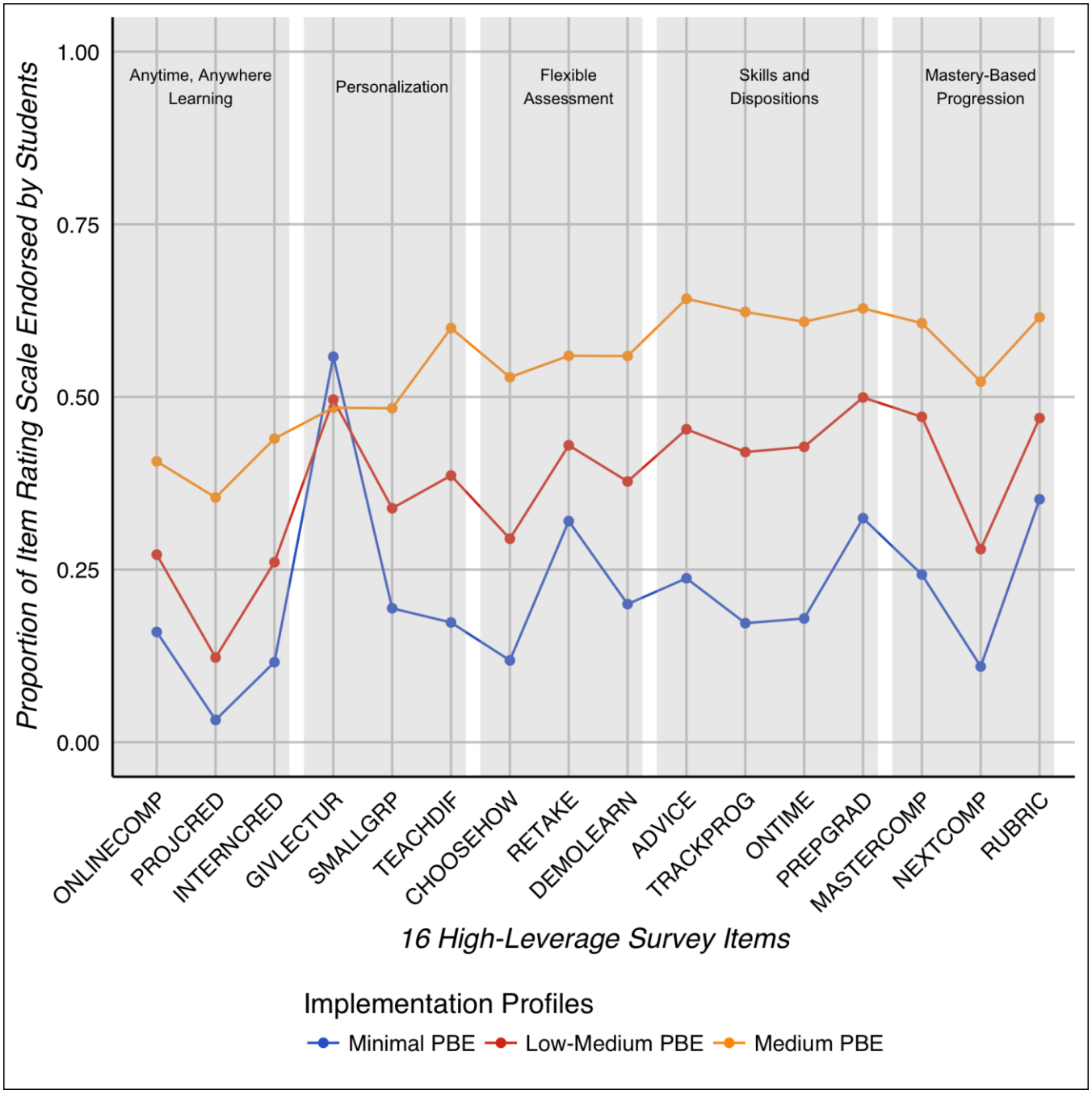


Figure 2 demonstrates that the pattern of students' responses to the survey items for the three exposure profiles was quite similar. The profiles differ from one another primarily with respect to level of exposure, which is visually apparent in that the profiles repeat the same general pattern of

responses but shift up or down the y-axis. For example, for each profile, the item related to completing course requirements online (ONLINECOMP) is about the same for each profile, relative to the other items within each profile. In other words, students in the Minimal exposure profile may have lower exposure to completing course requirements online than students in the other two profiles, but their level of exposure is consistent relative to their level of exposure to other items for that profile.

Consequently, we have labeled these profiles as Minimal, Low-Medium, and Medium to reflect the position of each of the three profiles. It should be noted that we have deliberately refrained from calling the profile containing students who reported the highest levels of exposure to PBE practices (i.e., the profile labeled Medium in Figure 1) the "High" group. This is to reflect the fact that no profile contained students who reported genuinely high levels of exposure to any of the PBE practices in the 16-item LPA. This can be seen in Figure 1 in that only a few items rise above the .50 mark on the y-axis. Importantly, these three profiles were present across all schools and within each school in similar proportions: Approximately 30% of students overall and in each school fell into the Minimal exposure profile; approximately 50% of students overall and in each school fell into the Low-Medium exposure profile; and approximately 20% fell into the Medium exposure profile.

There are a few notable variations from the similar pattern of exposure that repeats in the three PBE exposure profiles. First, regardless of profile, students' exposure to whole-class lecture (GIVELECTUR) is essentially the same. In other words, regardless of the profile to which students belong, they report experiencing similar amounts of classroom lecture, and this exposure is relatively high compared with students' reported exposure to other practices. This finding diverges from the pattern in which the Minimal profile has less exposure to an item than the Low-Medium profile, which in turn has less exposure than the Medium profile.

Second, when comparing the pattern of students' responses in the Minimal and Low-Medium with those in the Medium profile, a reduction in variation from item to item is observable. In other words, students who fall into the Medium exposure profile appear to have more similar exposure to different aspects of PBE, particularly in the flexible assessment domain, than their peers in the other profiles, whose exposure to different aspects of PBE seems more varied. For example, students in the Minimal exposure profile appear to have more exposure to the opportunity to retake assessments (RETAKE) than they do to the other two items in this domain (CHOOSEHOW and DEMOLEARN). This variation flattens out for those in the Medium profile, who appear to have similar exposure to all three items in this domain.

Research Question 1A: Characteristics of exposure profiles. Although students in all three PBE exposure profiles report some exposure to aspects of proficiency-based education, the results across all three profiles suggest that students' classroom experiences are still fairly traditional

Here we employ both the specific domains within the LPA (mastery-based progression, flexible assessment, etc.) and the data from the qualitative cases to examine the characteristics of the exposure profiles. As described above, for the most part, the LPA revealed three distinct exposure profiles that varied primarily in degree and were present in all schools. As a result, the majority of the discussion below considers the pattern of exposure that is common across all three profiles. In other words, our findings primarily discuss the overall pattern of exposure to various practices rather than comparing one profile with another.

Mastery-Based Progression: Although the schools have adopted the concept that students must demonstrate proficiency, opportunities for students to progress at their own pace are rare

Mastery-Based Progression

The three items that compose this domain are MASTERCOMP, NEXTCOMP, and RUBRIC. MASTERCOMP refers to the extent to which students must show teachers that they have mastered a competency (or proficiency) before they can move on, and NEXTCOMP refers to the extent to which students have the opportunity to move on when they are ready even if other students in the course are not. RUBRIC refers to the frequency with which teachers employ rubrics to assess student learning.

Students in all three profiles reported having higher exposure to the need to show teachers that they have mastered a standard before they can move on than exposure to the opportunity to move on when they have demonstrated proficiency on the standard even if others are not ready. At first, this result appears confusing because these two items are quite similar. However, in fact, they represent an important distinction. Although students must demonstrate proficiency to move on, if they do so, but other students have not, the students who have demonstrated proficiency cannot move on until the majority of the class has demonstrate proficiency as well.

This distinction between the expectation that a student will demonstrate proficiency to move on versus being able to move on to the next standard when he/she has demonstrated proficiency regardless of

where others are in the class was reflected in our interviews in the three sites. For example, in one school, both teachers and students referenced the term "teacher pace." Teacher pace refers to an expected pace of learning at which teachers are teaching the content and at which students should be progressing. While those students who are not up to teacher pace must keep working to get to teacher pace, those ahead of teacher pace may find themselves waiting for the majority to be ready to move on. This is not surprising if we consider the challenges, present in a traditional classroom, associated with moving students ahead of their peers. To do so requires more significant structural changes, such as flexibility in students' schedules and classroom structures that allow students to work on different content within the same class session, as well as considerable planning by the teacher to have content ready for students at different paces. In light of the fact that students, regardless of profile, reported relatively high exposure to whole-class lectures in their classes, it is not surprising that we heard only a few examples of the structures that would permit this kind of flexible classroom.

One parent observed the challenge she imagines teachers face when they attempt to implement mastery-based progression in their classroom:

> *I think in theory it's a good idea, having kids move on if they've already met that standard, but it's a lot to ask from one teacher. If we're thinking of a teacher in their traditional role where they stand up in front a class and teach the entire class one lesson. One kid doesn't get it, then you can't really have those other kids move on or else that kid is gonna be left behind.*

However, the same 9th grade math teacher reported how he approaches mastery-based progression in his course. He described planning his curriculum over a few days in the summer, using an online learning system in which he incorporated learning "playlists" or a series of lessons, in sequential order. Students enter his class, open their laptops, and begin wherever they left off in the last period. He explained:

> *I load everything they need on a playlist. I actually took two days this summer as workshop days individually and came in. My entire year is set up. Right now if you log in to my classroom, you can go from now until June and all of the assignments, and all of the modules that we're going through are there, in order. There's no dates assigned, but I do give my students on the first day they come in an outline that basically breaks up the year into six chunks, because that's how many modules we have. They know "Okay, this is how much time I really have to work on this." That's not hard, set dates, because if a student is finished early they can move on, and if a student needs to work on some standards they will keep working on that. I do keep a teacher pace. So Monday morning comes around, I will introduce the topic for the week or whichever day we*

> *finish a topic I will introduce the next one, do some example problems on the board, [and] go through it with the students.*

When asked what he will do with students who are ready to move on to the next course before the end of the school year, he explained that he anticipated that they would simply move to the next year's content when they are ready (which is already on another teacher's classroom playlist), without actually leaving his classroom.

In another school, a teacher questioned the traditional school schedule's capacity to meet the demands of a PBE model. Specifically, she wanted to experiment with a special class for students who had made it to 12th grade without demonstrating proficiency on the specific standards in science that would have appeared in earlier classes. Rather than requiring these students to retake whole courses because the students had not met proficiency on some aspects of the Next Generation Science Standards (which include both content- and practice-related standards), she proposed using a special course that could be tailored to the students and their areas of need. However, she explained that her proposal was rejected because of schedule and budget constraints.

Flexible Assessment: Assessment practices remain relatively traditional, but schools have increased students' opportunities to retake assessments to meet proficiency

Flexible Assessment

The three items that make up the flexible assessment domain are CHOOSEHOW, RETAKE, and DEMOLEARN. Two of the items (CHOOSEHOW and DEMOLEARN) reflect the idea that students can demonstrate learning in various ways rather than only with a paper and pencil assessment, and RETAKE refers to students' ability to retake an assessment if they did poorly the first time.

Students in all three profiles had little exposure to opportunities to choose how they demonstrate their learning or to demonstrate their learning in a variety of ways, but overall high exposure to the ability to redo or retake assignments if they do poorly the first time.

As reflected by the lower exposure to choice and variation in how students demonstrate learning, interviews indicated that assessment practices tend to be fairly traditional. Teachers reported that they mostly employ paper and pencil or computer-based tests and tend not to implement a lot of project-based learning, portfolios, or performance-based assessment. However, the schools have worked over the last few years to change policies about retaking assessments. The philosophy behind this change is the increased focus, in the context of PBE, on getting all students to know and understand the core content, even if that means they have to retake assessments to meet proficiency. As one principal observed:

> *I would say we're still fairly traditional in the way that we assess. A lot of our assessments are traditional end of unit type sit-down tests, whether they take them on the computer or they do them with paper and pencil. We do some projects from time to time, but that's really not common at this point, meaning a student has multiple ways to show you what they've learned. Now, if a student scores poorly on the test it's fairly common for a teacher to say, "Well, come and sit with me and we can talk about this, if you want to show me in a different way." I know that's happening from time to time in our school, but in general I'd say our assessment practices are traditional.*

In the context of the move to PBE, some schools and teachers have also made changes in the assignment and evaluation of homework. The Fields superintendent explained that the faculty and administration determined that grading homework did not fit in with the goals of PBE, which he explained is designed to help students improve:

> *We decided that homework assignments were no longer going to be graded. That was one thing that we determined, that we talked about how homework is about practice to*

help students improve. If we're going to continue to grade those traditionally and let those dilute what a student … If we're going to start grading things just for the sake of grading things and come out with a formula to deem a student proficient … we're doing the same thing, just in a different way.

Personalization: Students report high exposure, regardless of profile, to whole-class lecture

Personalization

The personalization domain includes three items: GIVELECTUR, which refers to students' exposure to primarily whole-class lecture; SMALLGRP, which refers to students' exposure to small groups and individual work with teachers; and TEACHDIF, which measures students' report of the frequency with which teachers teach the material in different ways to help students learn.

Across all three profiles, students reported frequent exposure to whole-class lecture, and this is the only item in the LPA analysis that does not vary by profile. In other words, regardless of whatever variation in exposure students may have to other practices such as flexible assessment or even other aspects of personalization, all students continue to experience the fairly traditional teacher-centered practice of whole-class lecture.

Students' exposure to small group work also presents interesting results. For the Minimal and Low-Medium group, students' exposure appeared to fall somewhere in the middle relative to other items in the LPA. However, the Medium exposure group appeared to have less exposure to group work relative to other PBE practices. This may suggest that the emphasis on demonstrating proficiency, as is the focus in Maine's PBE policy, may drive teachers to focus more on individualized instruction, which in turn may lead them to implement less group work. As the Fields math teacher who, as described above, planned out his whole curriculum as a series of playlists, explained:

I deal with group work a little differently. The MVP curriculum is designed as group work, and a lot of my stuff is designed as group work, but students can do it individually. I do leave that up to them. I've never forced groups, I don't assign groups. They self-select groups of up to three or work on it individually. That is their choice and their prerogative.

Although this teacher allows students to self-select into group work, the way he has designed his curriculum so that students can move individually, and at their own pace, might serve as a disincentive to working collaboratively.

It is also worth noting that the Minimal exposure group reported having less exposure to teachers teaching content in different ways than to small group work, reversing the pattern of the other two groups, for whom there was less group work but more exposure to teachers teaching content in different ways. This may suggest that the higher exposure a student has to personalized learning and mastery-based progression, the more that student has exposure to teachers teaching content in different ways but less exposure to small group and collaborative work. In other words, there may be a tradeoff, either real or perceived, between teaching content in different, personalized ways and conducting group-based or collaborative work.

Skills and Dispositions: Students may not perceive teacher support as necessary for them to be ready for life after high school, yet schools are assessing students on relevant skills and dispositions related to readiness

Skills and Dispositions

Four items compose the Skills and Dispositions domain: ADVICE, TRACKPROG, ONTIME, and PREPGRAD. Three items are quite similar in their focus on teachers providing students with support to take ownership of their own learning. ADVICE refers to the teacher providing advice and strategies to help a student to learn something; TRACKPROG refers to teachers showing students how to keep track of their progress; and ONTIME addresses teachers showing students strategies to make sure they complete assignments on time. PREPGRAD is a little different from the others in that it refers to students knowing what steps they need to take in high school to be prepared for what they will do after they graduate.

Perhaps not surprisingly, it appears that within each profile, students report having fairly similar exposure to all items on this scale. In other words, it appears that students who report that teachers provide advice on how to learn content also report that teachers provide guidance about how to complete assignments on time and how to track progress. One notable variation in this domain is that students in the Low profile, and to a lesser extent those in the Low-Medium profile, report lower exposure to the items about what their teachers are doing to support them than to the item referencing their own knowledge of what they need to do to be prepared for life after high school. Perhaps this variation suggests that students' sense of preparedness for life after high school is not so dependent on teacher support. However, students in the Low exposure profile do report less preparedness than peers in the other two profiles, so this item performs differently than the lecture

item in which there is almost no variation by profile. Yet it does raise questions about the importance that students place on teachers' supporting them to take more ownership for their learning as a key aspect of preparedness for life after high school.

In an effort to separate students' content knowledge from their development of specific skills and dispositions such as ownership and responsibility for their own learning, some schools have initiated separate grades associated with these skills. These "Habits of Work" (HOW) scores relate to students' behavior and to whether they meet requirements such as participating in class, completing homework, and coming to class prepared.

For example, Princeton began assessing "HOW" scores in the 2017/18 school year. There are four HOWs: preparedness, classroom engagement, effort and perseverance, and respectful conduct. Students receive a rating of 1–4 on each of these items every two weeks. The scores are included on the report card but are not factored into a student's grade. HOW scores are tied to athletic eligibility—students have to maintain good HOW scores to participate. As one Princeton teacher explained:

> *I think the real difference with habits of work is just making the students aware that we're looking for these. We're looking at them and seeing some basic things. Are they coming prepared? Are they doing their homework? Are they polite and respectful? … I think it's one of those things we've always looked at but we never maybe communicated as well, and I think habits of work kind of forces us to do it…. If a student came to me and they said, "Well I had a low habit of work grade"—usually, if that's low, then they're not doing their homework, and then their test scores are low. So it gives us kind of a focal point to start saying, "Well you're not prepared for class. You're not doing your homework. Is it time management? Is it that you need to come in and ask for help?" Now they're paying attention, and I'm paying attention more to the fact that they haven't been in, or the fact that they haven't had their book for a week, or what have you.*

At the root of the issue for all the case study schools is the goal to separate students' knowledge and skills related to the content standards from their behaviors, and to promote certain behaviors and values as part of the school culture and expectations for all students. One teacher explained:

> *We also are trying, we're really working on our culture and how the habits of work are also, what do we call them, common core values that we should all have. That will help us in all our life, not just at school … It should not just be about playing sports. It should be about our core values. We value responsibility. We value integrity. We value engaged people.*

Anytime, Anywhere Learning: Online and CTE courses are available to students, particularly in the 11th and 12th grades, but outside projects and internships are rare

Anytime Anywhere Learning

The three items that compose the Anytime Anywhere domain are: PROJCRED, which relates to the opportunity to earn credit for projects outside of school; ONLINECOMP, related to completing some or all of course assignments online; and INTERNCRED, related to receiving credit for internships.

For each of the profiles, anytime, anywhere learning was the domain for which students reported the lowest overall exposure. Specifically, students reported the overall lowest exposure to opportunities to earn credit for projects outside of school. Students also reported low exposure to the opportunity to receive credit for internships and somewhat more exposure to the opportunity to complete some or all coursework online.

This is in keeping with the results from the interviews. When asked about opportunities to engage in learning outside the traditional classroom, interview respondents tended first to reference online options, which in some cases many 11th and 12th graders are choosing. In fact, at Walden, many upperclassmen take several courses online, because of limited options in content areas such as foreign language. In addition, beyond online learning, many 11th and 12th grade students in the study schools attend the career and technical education center (CTE center), where they take courses and gain credentials in areas such as commercial driving, health, or childhood education. Although these are not internships in the traditional sense, the students have the opportunity to learn a skill, and many are able to practice the skill through some type of experiential component. Interestingly, according to the Fields superintendent, they have begun to explore internship options but have to be certain not to overlap with the content offered by the CTE center:

> *The CTE center offers building trades, offers automotive, offers welding, early childhood education, advanced technologies, and health occupations. As long as our internships do not provide something that can be offered at the CTE center, we're fine. So, we've had some internships with diesel mechanic type things. That's not offered at the CTE center, so it's acceptable for an internship.*

Respondents also indicated that there are limited internships available, in part because of the location of the schools—primarily rural schools with little access to businesses or organizations where students might intern. In addition, transportation issues can be a barrier to setting up and securing internships. As the Walden superintendent explained:

> *Well, we only have so many opportunities here. We're…small, rural…internships at the high school level are not easy to come by for anybody. It's just a lack of opportunities. We're trying to get more. We do have a garden project, started just this spring. Hopefully that'll give them more opportunity to be out of the building into another situation, but I think we need more of those types of opportunities.*

Although the Fields superintendent did describe some nascent efforts, our interviews indicated that the idea of anytime anywhere learning has not taken hold as a key aspect of the PBE reform effort in the three schools, at least not at this point.

Research Question 1B: Characteristics predicting membership. Student Grade Level and IEP Status Predict Profile Membership

After identifying exposure profiles and assigning students to profiles based on their greatest probability of membership, we used multinomial regression to examine whether student characteristics predicted students' membership in the profiles. Results from the regression model are provided in Table 4.

Table 4. Multinomial regression results for the model predicting students' membership in PBE exposure profiles.

Parameter	Low-medium vs. Minimal PBE				Medium vs. Minimal PBE			
	Est.	SE	p	Pseudo Rate	Est.	SE	p	Pseudo Rate
Intercept	0.82*	0.18	0	0.2	-0.54*	0.21	0.01	-0.14
Free and Reduced-Price Lunch	0.19	0.2	0.352	0.05	0.38	0.23	0.1	0.09
IEP	0.7*	0.33	0.033	0.17	1.78*	0.38	0	0.44
Unexcused Absences	-0.08*	0.03	0.007	-0.02	-0.08	0.07	0.27	-0.02
Unweighted GPA	-0.07	0.06	0.259	-0.02	-0.16	0.09	0.08	-0.04
Male vs. Female	0.43*	0.19	0.027	0.11	0.23	0.26	0.38	0.06
Gender Refused to Answer vs. Female	-0.1	0.53	0.853	-0.02	0.21	0.59	0.72	0.05
Grade 10 vs. Grade 9	-0.61*	0.24	0.011	-0.15	-0.34	0.23	0.14	-0.09
Grade 11 vs. Grade 9	-1.11*	0.31	0	-0.28	-1.13	0.2*	0	-0.28
Grade 12 vs. Grade 9	-0.72*	0.19	0	-0.18	-0.67	0.13*	0	-0.17
Race Non-White	-0.08	0.35	0.827	-0.02	-0.13	0.25	0.61	-0.03

Note. * = p < .05. N = 1823. Pseudo rates are calculated by dividing the estimates by 4 (Gelman & Hill, 2006).

The reference group for the analyses is the Minimal PBE profile. Results for predictors of membership in the Low-Medium and Medium PBE profiles are in contrast to being a member of the Minimal profile. The first set of results to note are the estimates that correspond to the various grade levels. These results show that being a member of grades 10, 11, and 12 are all associated with a significant decrease in the likelihood[6] of being a member of the Low-Medium profile or of the Medium profile compared with the likelihood of being in the Minimal profile, holding all other variables in the model constant. However, in the case of the Medium profile, only the results for grade 11 and grade 12 are significant. These results suggest that students in higher grades are less likely to report higher levels of exposure to PBE practices in their schools. This is logical in light of what we know about implementation. Because the state required 2020/21 to be the first class to graduate with a proficiency-based diploma, 9th grade students in 2017/18 are the first class of students for whom the changes that the schools have been making are high stakes insofar as they will ultimately lead to a different type of diploma. Of course, schools have been building toward this moment and toward implementing various aspects of PBE, as we described in the prior section, over the last few years, and have been implementing some aspects of PBE, especially in the 9th and 10th grades. Yet, the fact the 2017/18 is the first high stakes implementation year undoubtedly contributes to the result that grade-level membership is associated with profile membership; 9th graders in the survey sample were more likely to be in the Medium exposure group.

Interestingly, and perhaps more surprisingly, there is also a significant result observed for students who have an IEP. These students are more likely to be members of the Low-Medium and Medium PBE profile than of the Minimal PBE profile. To give a sense of the magnitude of the effect, we have calculated pseudo rates for these estimates. These results show that students with an IEP are 17% more likely to be members of the Low-Medium profile and 44% more likely to be members of the Medium profile than of the Minimal PBE profile. Both of these results are statistically significant. Overall, the results suggest that students with IEPs experience higher exposure to PBE practices in their schools. Although we do not know definitively why this pattern occurred, we can speculate that some practices associated with PBE, such as flexible assessment practices or efforts to teach students content in different ways, are also practices that students with IEPs may experience simply by virtue of having an IEP.

[6] Likelihood as per the multinomial log odds.

Research Question 2: Exposure and outcomes. Membership in PBE profiles is positively associated with students' self-reported engagement and negatively associated with their SAT scores

We next examined the relationship between students' PBE exposure profile and their engagement and SAT outcomes. We first present the results for the student engagement scores, followed by the results for SAT scores.

Students' exposure to PBE practices is positively associated with their self-reported engagement

We fit a 2-level multilevel model, with students nested in schools, using varying intercepts that allowed for variation in average student engagement scores across the 11 schools in our sample. Details of the analytic procedure and associated technical results are provided in Appendix D. Results that follow primarily concern understanding whether students' membership in schools was associated with their self-reported engagement scores, and whether a positive association existed between students' membership in PBE exposure profiles and their engagement scores.

First, the school that students attend did not appear to be associated with their level of engagement. The intraclass correlation (ICC) for the null model was .03. This result indicated that correlations between students' engagement scores within schools was minimal. In other words, differences in students' engagement scores was probably not associated with being a student in a particular school. Appendix D provides a figure for the distribution of students' scores across all schools, which can help in interpreting the ICC statistic.

Second, students' profile membership (in the Minimal, Low-Medium, and Medium profiles) does appear to be associated with student engagement. The model with the lowest BIC value included the binary indicator variables representing student membership in PBE profiles, a quadratic term for GPA and other student characteristic variables. A table containing the fit statistics for the sequence of model testing that led to this model is reported in Table D1 of Appendix D.

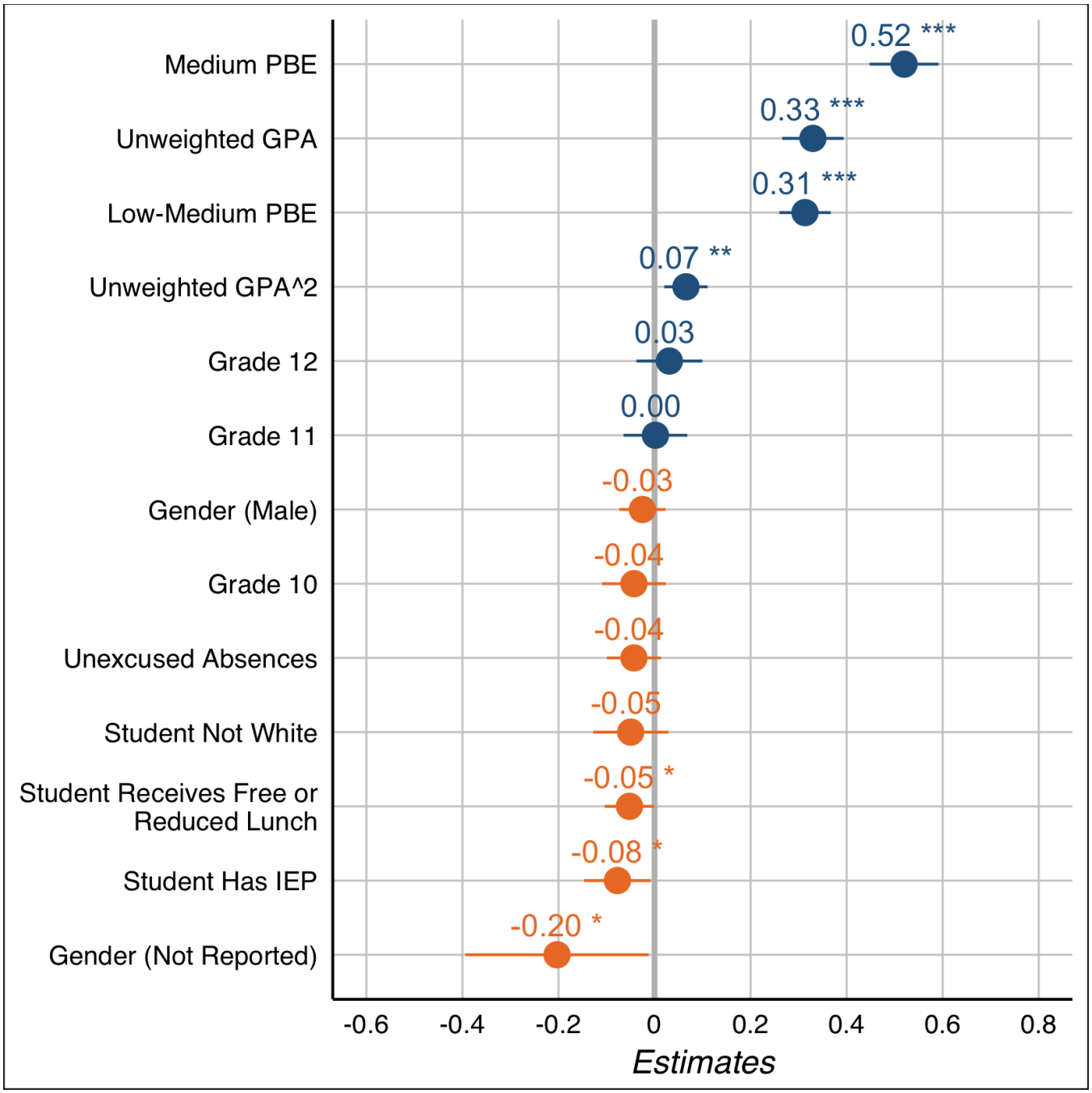
Figure 3 shows a graphical display of the coefficient estimates for this model (i.e., Model 4).[7] Detailed information for the regression estimates are provided in Table D2 of Appendix D. In the figure, we are

[7] In the figure, the location of the point estimate on either side of zero represents the magnitude of the effect, with values further to the right along the x-axis representing stronger positive associations between a predictor and the outcome and values further to the left on the x-axis representing stronger negative associations between variables. The relative magnitude of effects

most concerned with the two estimates labeled "Low-Medium PBE" and "Medium PBE." The estimate for "Medium PBE" is 0.52, indicating that students in this PBE exposure profile had higher engagement scores, on average, than their peers in the Minimal exposure profile. More specifically, students in the Medium exposure profile had engagement scores that are on average .52 standard deviations larger than those of their peers in the Minimal exposure profile, holding all other variables in the model constant. The estimate labeled "Low-Medium PBE" shows a similar pattern. These students had predicted engagement scores that are, on average, .31 standard deviations larger than those of their peers in the Minimal exposure profile. Further, the relative magnitude of each of these variables is among the largest effects in the model. Both of these estimates have 95% confidence intervals that exclude zero, which indicates that these estimates are statistically significant. The interpretation for coefficients for the numeric variables in the model (i.e., unweighted GPA and unexcused absences) corresponds to two-standard-deviation changes (roughly, from the low to the high end) of each numeric input. Overall, the results suggest that students who reported increased levels of exposure to PBE practices also had higher engagement scores than their peers in the Minimal profile, and these results do not appear to be driven by what school the students attend.

can be compared directly, and the estimates in the plot have been sorted in order of magnitude to make the comparison between effects straightforward. The bands extending from point estimates represent 95% confidence intervals. Estimates that have 95% confidence intervals that do not contain a value of 0 are significant at $p < .05$.

Figure 3. Coefficient plot for model predicting students' engagement scores



For the Low-Medium exposure profile, students' exposure to PBE is associated with lower SAT scores in 11th grade

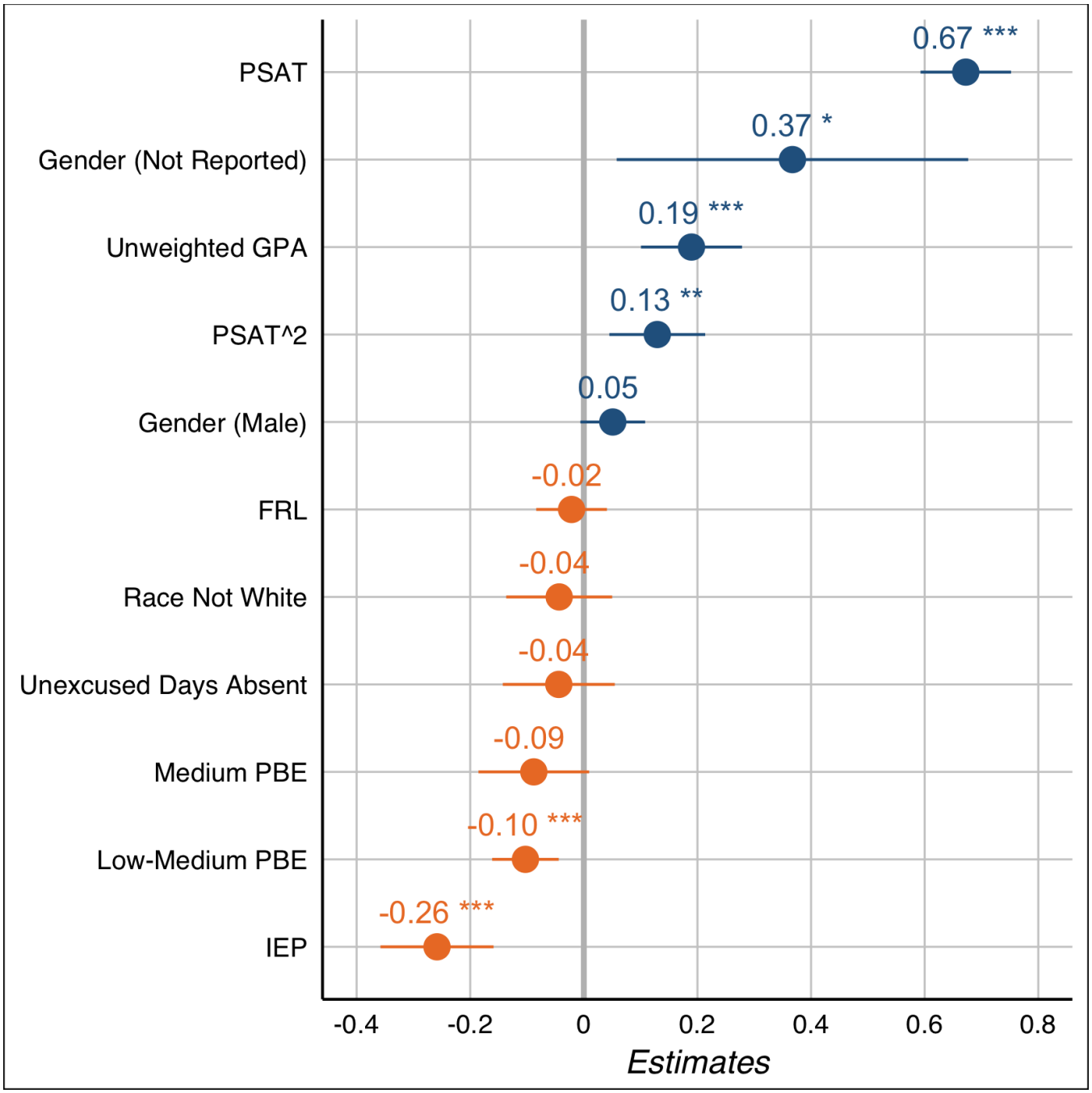
Similar to the models for students' engagement scores, we fit a 2-level multilevel model, with students nested in schools, using varying intercepts that allowed for variation in average students' SAT scores in

grade 11 across the 11 schools in our sample. Again, the primary goals of the analyses presented are to establish whether students' membership in schools was associated with their SAT scores and whether there was a positive association between students' membership in PBE profiles and students' SAT scores.

Like the model for students' engagement scores, students' membership in schools did not appear to be associated with their SAT scores. The intraclass correlation for the null model was .01, which indicates that the correlation of students' SAT scores within schools was minimal. Appendix D provides a figure for the distribution of students' scores across all schools, which can be helpful for interpreting the ICC statistic. Next, the inclusion of the binary variables for students' membership into particular PBE profiles was the most probable model given the data. This model also included a quadratic term for students' PSAT scores and for other student characteristic data. The one statistically significant result was that students' profile membership into the Low-Medium profile was negatively associated with 11th grade students' SAT scores. The estimate for the students in the Medium profile was similar in direction and magnitude but not statistically significant. Model fit statistics and estimates from the sequence of model testing are provided in Tables D3 and D4 of Appendix D.

The coefficient plot for the final model is provided in Figure 4. Again, we look at the two estimates that correspond to students' membership in specific PBE exposure profiles. These are labeled "Low-Medium PBE" and "Medium PBE," respectively. Both the estimates for exposure profiles are negative. This indicates that students who are a member of either profile have, on average, lower predicted SAT scores than their peers in the Minimal PBE profile. More specifically, students in the Low-Medium PBE profile had predicted SAT scores that were, on average, .10 standard deviation lower than those of their peers in the Minimal PBE profile. The result is statistically significant because it contains 95% confidence intervals that exclude 0 in the figure. Estimates for students in the Medium PBE profile are similar, because these students have SAT scores that are, on average, .09 standard deviations lower than those of their peers in the Minimal PBE profile. However, in contrast to the students in the Low-Medium PBE profile, the confidence intervals for this group include 0, meaning the estimate is not statistically significant. The magnitude of each result can be compared directly with other estimates in the model, showing that the relative magnitude of these effects is exceeded in size by several other estimates in the model. Both of these results are adjusted for students' gender, race, grade level, FRL status, and IEP status, and for the number of unexcused absences, unweighted GPA, PSAT scores, and a quadratic term for PSAT. Overall, these results suggest that students who self-report higher exposure to PBE practices have lower predicted SAT scores.

Figure 4. Coefficient plot for model predicting 11th grade student SAT scores



Research Question 2a: Student characteristics moderating outcomes. The data do not support that student characteristics have a moderating effect on student outcomes by PBE profile

We built on the 2-level multilevel models examined as part of Research Question 2 through the specification of interactions between the binary indicator variables for the PBE profiles (i.e., Low-Medium PBE and Medium PBE) and the student characteristic variables. Specifically, interactions were specified between these indicator variables and student characteristic variables that represented their gender, race, grade level, FRL status, and IEP status. Interactions were examined individually. Overall, no interactions were statistically significant in the models for student engagement and SAT outcomes. Results suggest that there is no evidence to support the premise that associations between membership in a PBE exposure profile and academic and engagement outcomes are moderated by student characteristics. For example, our analyses did not support the notion that a student in the Medium profile who has an IEP would have SAT scores that were lower than those of his peers in the Medium profile who do not have IEPs, as a result of his being a student with an IEP.

Research Question 3: Factors related to implementation. Implementation of PBE to date has focused primarily on changing grading rather than instructional practices, and the change has been met with some discomfort and uncertainty

The purpose of this section is to characterize the status of implementation and attitudes and beliefs about the PBE reform, across and within the three case study schools. Table 5 provides a summary of each of the three schools' demographics.

Table 5. Summary statistics for schools that participated in the qualitative component of the study

Characteristic	Princeton	Fields	Walden
Number of Students	545	272	139
Male	46.88%	48.37%	53.24%
White	90.28%	96.69%	95.68%
FRPL	41.10%	39.71%	56.83%
IEP	18.72%	18.75%	15.83%
Graduated	100%	97.44%	100%

Characteristic	Princeton	Fields	Walden
12th Grade	22.94%	28.68%	23.74%
11th Grade	24.04%	23.90%	27.34%
10th Grade	26.42%	24.26%	19.42%
9th Grade	26.61%	23.16%	29.50%
Unexcused Days Absent	11.88	4.62	4.26
SAT	984.69	997.74	946.46
PSAT	1027.98	908.08	874.59
Unweighted GPA	81.26	87.55	81.64
Engagement Total Average	42.87	40.74	42.78

In all three schools, the majority of the reform effort to date has focused on selecting the standards and implementing the new grading system. In addition, in two of the three schools they have implemented some formal rating of students' responsibility and engagement, in the form of a "habits of work" score, to evaluate whether students are turning in assignments, behaving properly in class, and showing respect for fellow classmates and for teachers. However, at the time of this writing, the schools had not implemented any significant structural changes, such as flexible scheduling, in support of PBE. In addition, although several respondents described the next step as one that would require shifts to instructional practices, this measure was not yet the focus of the reform effort in the 2017/18 school year. It is worth noting that given the potential changes to state policy concerning implementation of proficiency-based diplomas, it is uncertain what the path forward will be for implementing PBE-related initiatives.

As Table 6 shows, the three distinct PBE profiles (Medium, Low-Medium, and Minimal) not only are present in all three schools, despite variation in implementation of aspects of PBE, but they are present in proportions similar to those seen in the overall survey sample and to one another. In fact, the variation in percentages between the case study schools and the sample, or among the case-study schools, is not statistically significant.

Table 6. Percentage of students in each PBE profile by school, compared with survey sample.

Parameter	Minimal PBE	Low-Medium PBE	Medium PBE
Survey Sample	30.83%	50.96%	18.21%
Princeton	32.63%	49.05%	18.32%
Fields	31.22%	51.22%	17.56%

Walden	27.03%	48.65%	24.32%

The implementation efforts have, thus far, focused more on selecting the standards than on changing instructional practices

Overall, to date, the schools have focused primarily on the selection of standards to serve as the basis for defining proficiency in each of the content areas. This focus on standards is not surprising given the language of the state policy, which makes a passing reference to instructional strategies but focuses on the implementation of a proficiency-based diploma. Perhaps as a result of this policy focus, the schools have spent as much as two years identifying graduation standards within specific disciplines or content areas.

The determination of the graduation standards has generally been a department-level decision rather than a school or district decision. As we learned from many of the teachers across the three schools, teachers worked within their academic departments to identify the standards they would expect students to master by graduation. Even within schools, how this process unfolded and the expectations about how many standards must be met, and when and how often students will meet them, varies. For example, in the same school, one department identified their content standards and aligned them with specific courses. Therefore, the students may have only one opportunity, within a single course, to demonstrate proficiency on that standard. In the same school, another department identified a smaller number of standards that appear across classes and grade levels. In the former scenario, students have one course to meet the standards and have to retake the course if they do not meet those standards, whereas in the latter scenario (within the same school), students encounter the standard more than once during their high school career and have opportunities to demonstrate proficiency at different times in their coursework in that discipline.

One respondent expressed frustration that the state had not provided more guidance about the selection of graduation standards. He argued that, had the state provided these standards to the high schools, the schools could have moved forward with implementation more quickly and focused on what instructional or other changes might be needed to realize PBE more thoroughly. He explained:

> *I think that the biggest challenge right now for every district in the state is, I really wish the state had said, "These are the graduation standards, these are the standards that everybody needs"… If the state had done that, [if] they had done their homework and given that to us, then we could have worked on rubrics, assessments, and really, you know, the actual nuts and bolts of teaching … instead of … spinning our wheels … we could be so much further ahead if they just say, "Everyone's doing this." We could have*

really been knocking the lights out of assessment and learning targets and everything. It would have been, I think, just better time spent by everybody.

The PBE grading processes vary significantly within and across schools and this has led to a great deal of confusion for teachers, students, and families

The schools have approached the phase in of a new system of grading in various ways. The three case study schools implemented the following approaches in the 2017/18 school year:

- Fields implemented PBE grading in several courses in 9th and 10th grade, so that students have only the new 1-4-point PBE grades for those courses. Other courses were graded traditionally (0-100). Eleventh and 12th graders still have traditional grading across all classes.
- Princeton 9th grade students, and students in some other courses at teachers' discretion, received both a 1-4-point grade on the graduation standards and a traditional 0-100% grade for each course. All this information is available on the learning management system, but it appears that students and families have to know the standards-based ratings are there to find them. What appears on the first screen is the traditional grade.
- Walden implemented PBE grading only with the 9th graders, and only in some courses. A teacher we interviewed in this school explained that he provided traditional grades on all assignments and, in completing the progress report, converted those grades to a 1-4 rating on the standards. Other teachers in the same school reported having moved to a 1-4 scale for all assignments and for the report card.

Not surprisingly with the focus on selecting standards and transitioning to a new grading system, a great deal of attention has been placed on the 1-4 rating scale. Various respondents provided definitions for the numbering system, but essentially the 1-4 scale is as follows:

1 = not meeting proficiency on the standard
2 = working on or approaching proficiency
3 = proficient
4 = beyond proficient

However, this scale has not been employed in the same way in the three schools or even in the same school. In some classrooms, teachers are employing .5s between each number to allow for more opportunity to differentiate students' performance. Some teachers only provide ratings of 1-3 as options and do not offer 4s. In some classrooms at the time of our site visit in early November, teachers had offered no rating high than a 2 on certain standards because their expectation was that students would not be proficient in certain standards so early in the school year. One principal also explained that a certain standard might be demonstrated, through a given assignment, at a "level two

level of knowledge" in the 10th grade and would appear at a higher level of knowledge in 11th or 12th grade. Therefore, the 10th grader could only receive a 2 for that standard.

All of this variation in how teachers interpret and assign these ratings has been confusing for students and families because these "grades" are still consequential; students receive a score on these standards on their progress and end-of-year reports, and these translate into "grades" on a report card. Therefore, a student who had an 85% average in the traditional grading system may now receive a series of 2s at the end of 9th grade. This may feel, to the student and her parents, like a decline in her performance. It is also challenging for teachers to determine how best to evaluate students who are developing toward a standard. As one teacher explained:

> *It does make … me stop and think if a student has a two, what grade should they have? Why do they have a two? And so I do stop and think about that, knowing that they're working towards mastery. It's difficult because a two when you're just starting out, if it's the first time you've done something, it's good, right? I need to be careful that I'm not giving a student a low score, a low percentile grade, if they're doing a good job at working towards this goal.*

For parents and students, the paradigm shift is difficult, and it leads to debate over the meaning of the numbers and what the numbers translate into in terms of doing well in school, class rank, and access to college and scholarships. In each of the schools, there has been some discussion about how to treat traditions such as class rankings, honor roll, and the position of valedictorian. These markers of academic success and hard work are ones that students and families have valued over time, and the schools have made attempts to retain these traditions while changing the criteria for awarding them. In one school, students can apply to be the "valedictorian" and give the graduation speech, but selection is not based on their GPA. In another school, honor roll has been retained, but it is determined based on whether students are demonstrating the habits of work that are expected of them, and not on their meeting the academic proficiencies for each content area. However, interviews with students and parents indicate that there is still a lot of concern about the change. As one Princeton parent explained:

> *The one thing [my son] stressed about was the GPA. "How does this 4 or 3 on a paper correlate to GPA and how much money is a school gonna give me?" He stressed about that because he worked hard to be where he is today in his class.*

In the description of implementation and attitudes and beliefs for each case study school, we will return to the PBE-grading and the concerns expressed by parents, students, and some teachers.

Schools employ "habits of work" scores to continue to capture and evaluate the extent to which students take responsibility for work in and out of the classroom

As described above, Princeton began assessing four HOW scores in the 2017/18 school year. There are four HOWs: preparedness; classroom engagement; effort and perseverance; and respectful conduct. Students receive a rating of 1-4 on each of these items every two weeks. The scores are included on the report card but are not factored into a student's grade. HOW scores are tied to athletic eligibility, so that students have to maintain good HOW scores in order to participate. In part, these HOWs have been established to counteract the move away from grading homework, that is described above.

Fields began reporting Habits of Work scores in 2016/17, and teachers got together to determine a periodic HOW score for all students. In 2017/18, each teacher was responsible for reporting his/her own HOWs for each student every two weeks. Students received a score of 1-3 in four different areas: respect, responsibility, honesty, and engaged/invested learning, with a rating of 3 meaning that students met the standard. HOW scores are linked to eligibility for co-curricular activities, but the school is also experimenting with other ways to motivate students to develop their HOWS, such as assigning preferred parking spaces, providing snacks, or awarding early dismissal to lunch as a reward for good scores. HOW scores also affect a student's honor roll status. Some respondents talked about the role that the HOW scores play in promoting good habits and a positive school culture, and perhaps even in counterbalancing the drop in attention to deadlines that some teachers speculated might accompany the move to PBE. As one teacher explained:

> *I do know we graduated some students who were proficiency-based for ELA and math, and for the longest time we struggled with deadlines, and now the HOW scores seem to kind of be pulling that in somewhat, but some of our students have developed some bad habits that are really going to plague them about deadlines. That's one thing about proficiency based that's kind of difficult to maneuver, especially when they go on to college.*

In Walden, teachers have identified what the habits of work and habits of mind are, but they have not yet established a way to assess students in terms of these habits. They are grappling, in particular, with how to evaluate a student across the curriculum if, for example, the student shows persistence in math but not in social studies. In 2017/18 teachers did not assess students' habits within each class, aside from scoring effort and participation, which is something they have done on a 1-4 scale and separately from academic achievement for some time.

The schools have established formal structures to support students' academic and social and emotional development

In keeping with a focus on personalization, in each of the case study schools, formal academic support mechanisms are in place to ensure that students who need additional instructional guidance are identified for this support and have a way to access it that goes beyond seeking out a teacher before or after school. In Fields, the Student Support Teams play the role of observing and identifying students for extra help, and the school has built in an extra period in the day for students to access the academic support they need.

Walden has a model similar to that of Fields in that there is a designated class period each day for students to access extra academic support from their teachers. Teachers can "tag" students to come to their room for the 35-minute block to receive some "re-teaching." The principal explained:

> *If a student did poorly on an assessment, and they're ready to retake that assessment, that teacher can tag them to come in during that time, so … I mean, every teacher can have 8 to 10 kids that are in there for three or four different reasons. Teachers will tag kids, like a group of kids may come to a teacher and say, "Will you tag the six of us? We want to come in and study for our assessment tomorrow." And so the teacher will tag that group and assign them to a table in the room, so they can do a group study of the work that they're gonna be assessed on for the next day.*

The Princeton principal described the following academic supports: "Academic Recovery," which is an afterschool program that provides support to students who have been identified as not meeting the standards; "Instructional Study Hall," which is staffed by a paraprofessional who can provide help during the school day; and summer school, which is offered to students to make up work they missed during the school year, and presumably, plays a credit recovery role for students. However, one of the Princeton teachers also described the academic support groups the school has established to keep track of 9th grade students to ensure that they do not go unsupported if they are struggling:

> *A group of teachers, we work together; we're each assigned 15 students and we meet two or three days a week. We have one day in which we go through all the kids, my 15 kids' grades. On Monday, I go through and I record their grades, and I look to see who's struggling. Then we meet to discuss as a group … so we meet and we talk about these kids, and we try to do what we can to help them, whether we have to assign them to academic recovery or a structured study hall, or to talk with their special ed teacher, if need be. That's through the week, and then on Thursdays or Fridays, we try to contact—*

do the best we can to contact home by email or phone to see if we can do something to help these kids.

In addition to offering academic supports, two of the schools employ an advisory program to provide social and emotional support as well as career or next steps guidance to students. Advisory is composed of 10-15 students who meet several days a week with their advisor, who is a teacher in the school. This teacher plays several roles in relation to the students, with a primary goal of keeping students connected to the school. For example, in Walden, which has had advisory for several years, the advisor is responsible for supporting a student's personal learning plan. According to the principal, this includes the students identifying samples of work they are proud of from across their coursework and creating a portfolio of the work. Community service and career exploration of various kinds also takes place through the advisory time. In addition, students in the 12th grade complete a senior project, which is facilitated by the advisor.

Fields initiated advisory in the 2017/18 school year with a focus on the "guiding principles" that are part of the Maine Learning Results. These include effective communication, self-direction and lifelong learning, creative problem solving, responsible and involved citizenship, and integrative and informed thinking.[8] Each advisor has 10-12 students for 30 minutes every day, and they focus on different topics such as career exploration, growth mindset, and relationship building with students.

As described above, the bulk of the PBE work to date in each of the schools has included a focus on selecting graduation standards, aligning grading to the standards and proficiency-based diploma requirement, and separating out the content grade from a behavior or performance grade. In addition, the schools have implemented several PBE-related practices such as establishing academic and social supports to increase the personalization of instruction. We now describe the specific status of implementation in each of the three schools and parents', students', teachers' and administrators attitudes and beliefs about PBE in each school.

Overall, parents and students expressed discomfort about the PBE reforms, yet there is some variation in implementation and in attitudes across each of the three schools

Across the three schools, our qualitative research uncovered an overall discomfort among families and students about the reform effort, as well as concerns about the impact of the changes on students' motivation and on their future opportunities. Most teachers we interviewed expressed more

[8] The Maine Department of Education lists these principles, as part of the Maine Learning Results, here: https://www.maine.gov/doe/proficiency/standards/maine-learning-results.html

enthusiasm for the potential of PBE, but all expressed some reservations or described challenges related to how best to implement the changes expected of them. A few teachers were quite skeptical of the reform, as reflected in the low ratings they assigned to their own level of support (see Table B1 in Appendix B). By contrast, administrators tended to be more positive and enthusiastic about the possibilities that PBE presented in terms of its potential to increase their capacity to meet the needs of all students. Here we describe the status of implementation and attitudes and beliefs of key constituents in each of the three case study schools.

Walden High School

Implementation

In 2014/15, Walden High School implemented PBE grading for that year's 9th graders, who graduated in 2018. Although the state shifted the timeline for implementation before the year began, the school had already laid the groundwork to make the shift, and they went forward with it in the fall of 2014. According to the principal, after "a tremendous amount of pushback from parents and students, and ultimately the board," the school reverted to traditional grading in the following school year. Those students just graduated in June 2018 with a transcript that reflected a traditional grading system; the school converted the 9th grade PBE-based grades into percentages for each class.

When we interviewed the superintendent, in June 2017, he described the plan for implementation of PBE in the coming 2017/18 school year:

> *Our plan is to implement again this fall but probably do some type of combined proficiency tracking system with some type of traditional grading, because that's probably all we can do and get some type of acceptance.*

According to the principal, in an effort to avoid the strong negative reaction the school experienced in the first implementation, they attempted to communicate more with families about the change. However, we did not see evidence that these efforts led to greater understanding of the reform or more support for it. Although the school started the academic year with an event for all 9th grade students and their families, the event was scheduled at the same time as a sporting event and the turnout was very low. At the time of our site visit, in November 2017, students in PBE classes (math, English, science, and social studies for 9th grade) had just received their first progress reports in the week prior to our visit and, because of difficulties with the learning management system, teachers each created their own progress reports and did not work from a common format. Students received a rating of 1-4 on the standards they had addressed in class so far that school year, but there was a lot of confusion about how to interpret the reports and what the highest number was that students could receive for a particular standard:

Parent: *And with the progress report it was still the same thing. They got 2's and 1's and you didn't know if it was a 2 out of 2, 2 out of 4, were they working on it? Had they finished the standard?*

Interviewer: *Was there information that came home? I think one of the teachers said that they wrote something in that like—*

Parent: *Only one teacher did.*

Interviewer: *Okay.*

Parent: *So one teacher wrote in and said we're working on this standard but for all the rest of them there was nothing.*

Interviewer: *Okay, it was just the … number—*

Parent: *Just the one teacher wrote that they were still working on the standard.*

According to one of the teachers we interviewed, she graded assignments with the proficiency-based score and explained what the highest possible score could be (which was either a 3 or a 4, depending on the assignment). She explained:

> *Then, if it was a 3, they [the students] always want to know how they can get a 4. Usually I come up with a project or an essay or something that will come after the assessment. That usually satisfies them, as long as they know that they can reach a 4, preferably before grades close. They accept that. Parents actually are a harder sell on that, where they're not in my classroom and get my whole spiel. We've had parent meetings in the evenings, which I haven't been able to attend. I've mainly talked to parents on an individual level, by email or phone. Again, it usually feels like I'm a little bit under attack because they say: "My kid only got a 2 on this, what's going on? She's never gotten a 70," because that's what we translate that to …*

Overall, there appeared to be some real confusion about the implementation of the new grading model. At least some of the teachers were still using the traditional 0-100 percent for assignments and homework and were making the conversion to a 1-4 rating for the class or for specific standards only for purposes of communicating in the mid- or end-of-semester report. A student offered her summary of PBE grading:

> *Well, in math we only have 0-100 still. But in science we'll have just proficiency. Then, in civics, he'll give us a proficiency 1-4, and then he'll translate it to an actual grade. It's confusing for us to know if we actually get it, or we don't get it, or we pass it, or we don't pass it.*

Based on our interviews, it appeared that teachers tended to employ traditional methods and assessments. Certainly students and teachers interact a great deal; some teachers have their students for several years. In addition, the advisory system at Walden, whereby students and teachers are together for 4 years, also provides another avenue for personalized learning and the development of social and emotional, or 21st century, skills. Although respondents did not speak at length about the senior project, that too seemed a ripe opportunity for supporting students in examining topics of interest to them and in driving their own learning. However, our interviews did not uncover a sense that the teachers and students saw these various aspects of the school experience as related to their transition to PBE. These practices, rather, seem to be running concurrent with the PBE work, which to date seems to have focused squarely on selecting the standards and transitioning to the new approach to grading student performance.

Attitudes and Beliefs

As described above, Walden implemented PBE in 2014/15 for just one year, receiving a strong negative response from the school community. The superintendent described a tight-knit community but one in which some members are leery of the reform. He explained that the more academically oriented students and their parents were the ones who were most resistant to the change, particularly because the new rating scale meant that a student who might have received a 98% average in a class and the student who might have received an 80% average in the same class in the traditional model would both be considered proficient and would both receive a "3." He explained that this did not sit well with some members of the community when they implemented the PBE grading system in 2014/15:

> *Maybe somebody whose traditional grade [was] a 98 was getting a 3 and not a 4, or they might have got a 3 and one of their peers who maybe used to get an 80 or an 85 also got a 3, [so] they're all getting a 3. They all met the standard. They were worried that their child wasn't excelling or shining; given a 3. Although they could get a 4 and I think we have to be real clear on the fact that there are students who want to excel can get better than a 3.*

In the parent focus group, parents expressed concerns not only that the new model did not distinguish sufficiently among students but also that their children would lose motivation to perform and did not feel inspired to work as hard as a result of shifting to the new grading system.

Students also raised concerns about PBE grading. They expressed overall confusion and frustration about receiving "2s" on their report cards for standards that the teacher indicated the students could not receive a 3 or a 4 on at this point in the year, because the course had not sufficiently covered the content. They argued that therefore, the standard and rating should not be on their report card and sent home to parents, if there is nothing the students can do to improve that rating. They also raised

concerns about whether the new PBE grading model sufficiently distinguishes those students who excel:

> *Most jobs will look at your grades and say: "Oh, you really excel in your grades and you're good at keeping them up. We're gonna hire you." Now it feels like they can just hire anybody because they're proficient in one standard.*

In addition to community concerns regarding the new grading model, some teachers expressed concerns about their capacity to teach in the ways that PBE demands. For these teachers, the concern was not about the grades but about how to meet all students' needs in the PBE model. One middle school teacher, who is also a parent, explained her concerns about implementing PBE:

> *The thing about proficiency-based education is that it doesn't just change the way we grade, it changes the way we teach. Which when I think about next year, and I'm trying this year to teach as though this is how I would be doing it next year, I feel like a first-year teacher again. I also think that they're requiring a lot … from their teachers, but we aren't getting any extra help. You know, put more Ed techs in the classroom. If I have students on all these standards all through ELA, which has tons of standards, I don't know how I'm going to do this, by myself, and do it well, and I am committed to my students. I want to make sure that I'm doing everything I can; I want them to enjoy being here, I want them to be learning, and I want to be sane while I'm doing it. And I just don't see how this is all going to come together without some major changes in the classroom in order to make it work. I worry. I'm worried.*

Another teacher, who has been at the school for more than 30 years, explained that he does not feel capable of implementing the instructional changes that a shift to PBE requires. He described visiting other schools that are implementing PBE and his ambivalence about teaching in the ways that PBE requires:

> *I visited schools that this group [of students] was over here doing this thing, and this one's ahead of that group, so they're in that corner doing their thing, and … you know what I mean? And this group's working on this standard, and that one's working on that standard, and this one's behind so they have chances to … I won't stay in this if it becomes that. I can't. I'm three years away from retirement, I can't … that's just … I don't have that skill. I don't have that.*

Although there was a fair amount of trepidation at Walden about how to implement aspects of PBE, particularly the new approach to grading, there were also teachers who expressed enthusiasm about

the reform because it aligns with their desires to teach in ways that might engage their students and inspire them to work hard and enjoy learning. As one teacher explained, "PBE is helping me find things in the curriculum … helping me think of good ways for the students to prove their proficiency, and beyond. It lends itself to having more creativity." This same teacher also observed a subtle shift in her students' thinking about their performance. She explained:

> *I am noticing their language is changing. Instead of asking me: "Is this wrong?" they're saying: "Can I go deeper with this, or is this already a 4 level?" Their language is changing, and I think it's for the better even though they might not be adjusting to it as well. That kind of mind shift is happening.*

Walden's path forward in terms of implementing PBE, with more content areas and grade levels, and with changes to instructional practices, was unclear. The faculty include both several teachers close to retirement and some new teachers who have been very active in the school's efforts to make the transition. However, the community's animus about the reform appears to be as strong as it was in the first attempt at implementation, in 2014. Now, with the changes afoot statewide, it's possible that the school will stop implementing the reform altogether.

Fields High School

Implementation

According to the school principal, in the 2017/18 school year, Fields High School implemented the PBE grading system for 9th and 10th grade students in their math, English, social studies, and science classes. However, students at all grade levels had some exposure to PBE, insofar as their teachers in math and English evaluated their proficiency on a set of standards that the teachers collectively identified as necessary for graduation. However, for the upperclassman (11th and 12th graders), the school converted these proficiency-based grades (1-4) into a standard 0-100-point grade for purposes of reporting. The principal and students additionally reported that the middle school has also provided students with some exposure to PBE so that, as these students move to the high school, they will not be surprised by the new approach.

The Fields superintendent, who has just completed his third year in district, described some of the shifts he has observed as teachers move toward PBE. Although he recognizes the role of grading as part of the reform effort, he discussed instructional practices that are connected to PBE and the need for the district to provide teachers with guidance and support to make some of these shifts. He explained that he witnessed a desire among some staff for a quick fix, or a program that would help them shift their current approach to evaluating students to the PBE grading model, something we saw many teachers talk about across interviews. He explained:

> *I think when PBE was first introduced, educators thought a particular computer program or student management system would be able to help us deliver the expectations of PBE. Now several years later, we understand that PBE is so much more than a resource or a particular practice. PBE is a complete shift in the way we set learning objectives, track and assess learning, and facilitate the learning process. Technology is just one small piece to the PBE puzzle.*

Although he asserted that grading is not what PBE is all about, he also explained that more work was needed to develop a common understanding of the grading scale and what proficiency means and a common system for recording this information. He worked with teachers over the last year or more to gain some shared understanding of what a 1, 2, 3, and 4 is on the new proficiency scale and to dissuade staff from thinking that 4s are unattainable or "unicorns," as he described them. In addition, he did select and establish a policy stating that all teachers would use the same learning management system to record proficiencies:

> *We had groups of teachers who have been piloting the Empower Program, while others have been tracking standards with a Google Document sheet created in the district. It was just all over the place. We then made the decision to start the next school year with the same expectations for reporting with Empower. That is our proficiency-based platform, so everyone was going to use it. For the first time ever, every learner in the district had a similar progress report, which looked the same and was based on the same proficiency scale. These are big changes to initiate, but they are necessary. We call it "Ripping Off the Bandaid!*

In addition to some of these shifts, another change in policy, as part of PBE implementation, was the decision to stop grading homework. In the 2017/18 school year, homework was no longer a contributor to the students' grades. Rather, to distinguish between what students know and can do related to course content versus how students show responsibility for their work, the district implemented Habits of Work scores as a way to capture and record students' responsibility. Students receive a Habits of Work score for each class every few weeks, and this is, in part, a reflection of whether they are completing homework. This score does not factor into their course grades but does serve as an eligibility requirement for participation in sports, band, theater, and other after-school activities and is the determining factor for earning honor roll status.

Other practices at the district that might be associated with PBE or student-centered learning include the initiation of advisory for all students this year, and the senior project, which has been around for several years. There is also some early discussion about allowing students to move on mid-year if they complete all the content for a course, particularly in math, where the student could continue within the

same physical classroom but advance in content. In addition, as described above, there are limitations to the types of internships the district can offer, but they have begun to implement some, such as giving students the opportunity to intern in the elementary school.

Attitudes and Beliefs

Although several teachers expressed enthusiasm and support for the shift to PBE and consistently spoke positively about the school leadership, respondents did also express concerns about PBE. One teacher, who described himself as a "10" on a scale of 1-10 in terms of level of support for PBE, explained why he thought the shift to PBE was good:

> ***Interviewer:*** *Do you see kids who are less academically inclined doing better in this model? Do you see that this benefits kids on IEPs or kids who've struggled—*
> ***Teacher:*** *Absolutely, because they were the ones that would fall between the cracks a lot of times and their goal most of the time was just to pass. A lot of times if they could come back and redo one test or a project or something just so their overall average came up to a 70, they were like, "Done with that one, okay." You don't see that anymore. I find that they, although at first it was like, "You're holding me accountable for everything? You mean I have to pass everything?" Well, it's not like necessarily you have to pass everything, but we want to have you at least past these standards. I find there is a greater sense of accomplishment for those kids. I feel there is a renewed sense of, I don't know, their own sense of accomplishment. I think they feel better about themselves. I think that's definitely a positive for everyone.*

This teacher, who expressed a generally positive attitude about the shift, also acknowledged the challenges. He explained that he wanted his own children to go through the school and experience PBE but that a lot still had to be "ironed out" about the new system.

Other teachers expressed concerns that the move away from homework and the more flexible approach to retaking assessments might have a negative impact on students in the future. As one teacher explained:

> *I really truly believe in proficiency based. I've seen the best work come out of students since I've been using it, out of all our students, even the ones who struggle. What my concern is, is preparing students for college. I think we've, we're working on the habits of work, which are helping us get there. I do know we graduated some students who were proficiency based for ELA and math, and for the longest time we struggled with deadlines, and now the HOW scores seem to kind of be pulling that in somewhat, but some of our students have developed some bad habits that are really going to plague*

them about deadlines. That's one thing about proficiency based that's kind of difficult to maneuver, especially when they go on to college.

This concern was expressed more vehemently in the student and family focus groups. Parents and students alike expressed concerns that some students were lazier or less motivated as a result of the shift to PBE. One parent, describing her daughter who works hard but struggles academically at times, explained that she both valued the shift and found fault with it:

She doesn't understand how come she studies for a test, studies for a test, and she can't pass it, she can't pass it. So I find with the proficiency based … I find that's helped her a lot. She gets that second chance. But then I also find that it's made her lazy. Like she'll think, like, "Well, you know, I'll be able to try again."

However, although some parents agreed with this concern about the decreased motivation, another parent explained that it really boiled down to the student-teacher relationship: "If they have a teacher who motivates them, who pulls things from them, who connects with them, then they want to please them. They want to do more, because they have that connection."

Students themselves also expressed the concern that the PBE model had negatively influenced their motivation and drive. One 10th grade student explained:

I just feel like I'm not getting challenged enough because I know if I don't pass it, I can just do it again and do it again … so knowing … that I did have a bunch of chances I would … and I think more kids have gotten kind of lazy, because they know they can just keep on doing it and kind of coast through.

However, the students also acknowledged that the PBE model can benefit students who are motivated and working hard but just need more time and practice to master certain skills. One student noted:

You can tell the kids who try and who don't, and the kids who try but just take longer, I think it's great for them, and I don't see anything wrong with it. It's just the kids who are able to do it and they don't because they just know that they can do it next time.

It is not clear how Fields will proceed with PBE implementation in light of the potential for a change to state policy. However, the superintendent has been a strong advocate for the reform in his three years at the district and has brought new resources to the district, including additional professional support and development opportunities beyond what the collaborative of districts offered. He has also worked to communicate with students and parents about the reform, holding a series of separate meetings

with students and with parents in each of the grade levels to talk about the changes. All this might suggest a strong commitment to continue to pursue the reform, regardless of changes to state policy.

Princeton High School

Implementation

In Princeton, implementation of PBE has been focused on selecting the standards and implementing the new grading system, primarily for 9th grade students in the 2017/18 school year. However, some teachers had begun implementing some aspects of PBE in years prior, and some teachers are implementing aspects of PBE, such as providing information on the standards and what they are measuring with a given assessment, in all their classes rather than just in 9th grade. In 2017/18, many 9th grade classes officially became proficiency-based, insofar as the teachers provide information about the standards being measured and assign proficiency-based grades in addition to the 0-100 grade. The principal explained, "We're just kind of dipping our toe into the water. We haven't jumped in 100% at this point. We've been talking about this and dabbling in it for a number of years now."

The school uses PowerSchool to report on student performance. All students have an account and can log on at any time and review their performance on all class assignments. Because the school has a hybrid system in place whereby students still receive 0-100 grades for their classes, the students must click on the grade to drill down into their performance on specific assignments or on specific standards required for the class. It is only after clicking a few levels down into the system that students see their 1-4 proficiency rating on the standards. According to the students, at this point, only some teachers are entering these ratings into PowerSchool for assignments and midterm grades.

Some teachers have moved toward PBE more than others, and the principal reports that the shift has been easier for some teachers. His goal is to not move so fast that he alienates committed long-term staff. He explained:

> *We can't move so fast that people are frustrated and feeling like they're not doing a good job or feeling that they're overwhelmed, but we can't also put up barriers to try to implement this system.*

Teachers also described efforts to ensure they are reaching all students in their classes. One teacher, who described himself as somewhat ambivalent about PBE, explained that he did believe it made him more attentive or "alert" to the needs of his students. He stated:

> *Well, I will say, since we've gone to this new format, I have done a lot more conferencing with the Algebra 1A class because maybe, I don't know, maybe I care a little bit more. Maybe I've changed my attitude towards it, but I think I want them to meet the standard,*

and the only way to meet some of those kids is to sit down one-on-one and to make sure that they're step-by-step, and I don't think the material's changed to personalize it or anything. I just think that I've become more—tried to get to know them on a math level. I try to get to know all the kids, you know what I mean? I mean, I just think I've done a lot more conferencing, and a lot more walking around the room, and a lot more helping them than I have before.

In addition to the selection of standards and hybrid grading practices, another activity that might be associated with PBE is the school's newly implemented effort to keep track of 9th graders. Each teacher on the 9th grade teacher team is assigned a group of students for whom he/she is responsible. These teachers monitor the 9th graders' performance, call home to report both challenges and successes, and make sure students have the supports they need. In addition, the district has a curriculum director who is responsible for supporting teachers in transitioning to PBE. Her role has been valuable in providing the support the principal explains is so critical as the school works with teachers to shift practices:

I wouldn't dare say we're making progress yet, because we just kind of started talking about instruction. And I think when you do that, you just ask them to try it with one unit. You don't say, "Throw out all your units and throw out all the things that work," but "Let's start with one unit that we can [make] student centered or proficiency based." But in order to do that, [you've] got to provide more examples, you got to show teachers how it's being done, you have to bring in people who are knowledgeable about student-centered learning, you need to provide teachers opportunities to see it in action. So, it's just … We kind of got the ball rolling a little bit a few weeks ago, but we've got to keep the ball rolling, and that's really a focus of mine, and [of] our curriculum director.

Attitudes and Beliefs

The school's hybrid approach to implementation of PBE grading seems to have assuaged some of the general anxiety for teachers, parents, and students. Some parents indicated that the focus on ensuring that students really learn the content and are not pushed too fast is something they value. One parent explained:

I think the benefit is that I know that when they move on to the next section that they actually understand the one that they just came out of. They don't come out of it with an 80 and move on, right, because that means they don't quite understand it, but they're moving on because everybody else is moving on. So that won't happen. So that's good, so that I know they're getting the most that they can out of each class.

However, parents and students at Princeton are still highly critical of PBE for the same reasons as they are in the other schools: a concern that the shift to the different way of evaluating students does not differentiate performance in the way the old system of class rank did, and concerns that the new approach reduces students' motivation to work hard. One parent explained:

> *My kids actually study and they'll work hard for their grades. So their concern is, "Mom, if I'm working this hard, but I'm gonna get the same grade as everybody else, who maybe isn't working hard … How am I gonna get into medical school?" was his question. "Mom, how do I get into medical school?" … And I'm like, well I don't know. I don't know how the colleges will then determine the children who have really strong work ethic and—you know what I mean? The child who pushes himself.*

Students reiterated these concerns but also noted some of the potential benefits of the move to PBE. For example, one student observed that PBE seems to focus the classes a bit more. He stated:

> *What I've noticed with this PBE is it's added more structure to the classroom for the upcoming grades because it kind of gives the teachers something to aim for. Before they were kind of just going through a book and going chapter by chapter, but now they have this list of goals they have to meet.*

The students also reflected on the difficulty of shifting the grading paradigm in the midst of their academic careers and pondered whether it would be better to implement PBE starting in the early grades rather than starting with high school. One student noted:

> *It's kinda hard to think of it in a different way just because you've grown up on the zero through 100 scale, so as soon as we get a number, we want to try to convert that into a 75 or a 90. I think that's the tough part. If you start at Pre-K level and they don't even think about that zero through 100 scale, I think it would be less … I don't know. It'd be a little bit easier to understand what a 3 means instead of looking at it as a 75.*

Teachers, in general, were supportive of PBE insofar as they appreciated the focus on ensuring that students learn important knowledge and skills before they move on. However, they raised concerns about their capacity to effectively implement the reform without adequate time and training. As one teacher explained:

> *I think it's made it very difficult for teachers. We've got, I don't know how many teachers we have in this building, 40 something, and you can only send so many of them to trainings, and on top of that, the time. Teachers don't know what they're doing, and if*

> *you can't give them more concrete examples and discussions about what they need to be doing, then it's just a mess.*

A special education teacher raised specific concerns about what demonstrating proficiency will look like for some of her students and her concern that the PBE reform efforts have not adequately considered the ways in which implementation of the reform must be modified for some of her students. She described her concern that Maine has yet to determine how to address meeting proficiency for students in special education classes:

> *I think that we are just all very confused on how do we get those kids with those big learning gaps, or with those learning disabilities to be able to meet those standards? So I do a lot of support for those kids, I have a lot of freshmen this year that are mainstreamed for freshman science and World Studies, and we're only till the end of quarter one, and they're already not able to meet the standards from quarter one of that class, so it's very concerning; how are we gonna do this? How is this going to work? And to be honest, nobody really … has a good answer for us.*

The approach of the leadership at Princeton to shifting to PBE has been gradual, with recognition of the need to not overwhelm staff, students, or families with the changes. It is unclear what the path forward will be for Princeton with the possible changes to state policy and, as a larger school, making structural changes could certainly be more of a challenge. However, this gradual approach to implementation may ultimately support the school in making the shift to PBE overtime.

Discussion

Our study of PBE in 11 schools in one state raises important questions about the process of implementation of student-centered, proficiency-based reforms. Of course, the study raises more questions than it answers, but these questions are critical for other schools, districts, and even states to consider.

Why do we see three distinct profiles of student exposure in every school, despite differing contexts and approaches to implementation?

We speculate about several different reasons that we see the three distinct PBE exposure profiles. First, the presence of these three profiles in every school and across all the schools in similar proportions, is likely in part because of the nature of student-centered, proficiency-based reform. The recent reform has been conceptualized to include several different, overlapping principles with related structures and instructional practices (JFF, 2014), some of which, such as demonstrating mastery on a set of standards, are more concretely part of the current reform effort in Maine, and some of which, such as employing a rubric or allowing retakes, are more broadly part of many schools' and teachers' practices. Therefore, when we study implementation of student-centered, proficiency-based education and understand this to be a broader reform effort and not one focused solely on a shift in grading practices and graduation requirements, we are probably going to uncover practices that predate this current movement and that are not explicitly associated with a school's most recent efforts to implement Maine's proficiency-based diploma requirement. Therefore, the presence of a range of practices related broadly to student-centered, proficiency-based education across the schools may have less to do with the status of each school's deliberate implementation of PBE and more to do with the nature of teaching today, where some practices are more common than others (e.g., anytime anywhere learning is less present than flexible assessment). In addition, the variation we see within schools can likely be attributed to the fact that teachers' practices vary and that not all students experience the same teachers; for these reasons, some students are going to experience more exposure to some practices than others.

However, other factors may also be contributing to these three profiles. As our results suggest, IEP students are more likely to report higher exposure to PBE practices. As in our observation above, we would argue that some practices that are conceptualized as part of student-centered, proficiency-based education also occur in contexts in which there is no explicit reform effort in place. Specifically, some practices associated with teaching students with IEPs mirror those associated with student-centered, proficiency-based education. Thus, it may be that what we might perceive as evidence of the implementation of PBE related to the recent reform effort could in fact result from exposure to

practices associated with meeting the needs of students on IEPs, such as differentiated instruction, flexible pacing and assessment, and so forth.

As this study set out to understand more about implementation of PBE specifically in Maine, and more generally, to add to the field's knowledge about how to measure implementation, the results suggest that when attempting to measure such a broad reform effort, it can be difficult to determine where the reform ends and other practices and structures begin. What we know is that in the schools in our study, the focus has been mostly on the selection of the standards and the transition to a proficiency-based model for grading. Thus, we see these efforts reflected in the LPA. However, some of the instructional practices that are conceptualized to be a part of student-centered, proficiency-based education (such as flexible assessment, or anytime, anywhere learning) are also occurring in some degree across the schools, suggesting that these practices may be present regardless of the explicit reform effort and its focus. This suggests the need for research that continues to examine the nature of implementation of such a wide-ranging reform effort, as well as the need for studies that focus on fidelity of implementation and outcomes related to more focused practices.

What student-centered, proficiency-based practices appear to be most challenging to implement?

Based on our quantitative analysis, we see that certain practices are more common and more widely employed across the three profiles. These include practices such as enabling students to retake assessments, using rubrics, and enabling students to progress through demonstration of mastery. The greater exposure to these practices may have to do with the state law, which focuses all schools on the end goal of a proficiency-based diploma, and some of these practices (such as mastery-based progression) are clearly related to that goal. However, there are other aspects of PBE to which students reported overall less exposure that might also be considered directly related to the proficiency-based diploma requirement. These include the domain of anytime, anywhere learning in addition to items that refer to students' opportunity to choose how they demonstrate their learning.

Furthermore, as discussed above, students report less exposure to the item that refers to students' ability to move on when they are ready than exposure to the item that refers to the expectation that students demonstrate they have mastered the competency/proficiency before they can move on, both of which appear tightly connected to the state law. This distinction, as described previously, is an important one for understanding what may be easier versus what is more challenging to implement. As we discussed previously, it appears to be much easier to establish the expectation that students cannot move on until they have demonstrated proficiency on the standard the class is studying than it is to develop the structures to support students to move on whenever they are ready; doing the latter would require the building of structures to support students who may finish all course content in March of a

school year, along with different structures for students who master some but not all of the course content by the end of the school year. Although the schools in the study have begun to explore or consider what they might do to address this kind of flexible pacing, they have yet to establish many structures that make it possible. The results related to teachers' use of class time for lecture are critical here as well. Regardless of what other practices students are exposed to and how frequently, they are all experiencing a similar, and significant, amount of whole-class lecture. This suggests that teachers' use of class time for lecture is perhaps the most entrenched instructional practice for teachers.

Do some students have more access to PBE practices? Do PBE practices have more importance for some students than for others?

Based on our analyses, IEP students, on average, have more exposure to PBE practices than peers who do not have IEPs, putting them in the Medium exposure profile. However, we cannot claim that students with IEPs receive any particular benefit from membership in the Medium exposure profile. We observed no statistically significant interactions that allow us to find that membership in the Medium or Low-Medium profile is associated with higher outcome scores for any given subgroup (such as IEP, FRL). In fact, we cannot make any claims about the value of PBE practices on engagement and academic outcomes for these subgroups.

However, we did observe a relationship between membership in the Medium and the Low-Medium profiles, and student engagement. This should not be interpreted as a causal relationship, and it is not one that varies by student subgroup. It is hard to determine why this positive correlation exists except to posit that students who are reporting more exposure to practices that personalize instruction, and allow for greater flexibility, are also students who feel more engaged in school.

In terms of academic outcomes, we do observe that membership in the Low-Medium exposure profile had a statistically significant, albeit small, relationship to SAT scores: Students in this profile had lower SAT scores than their peers in the Minimal profile. The size and direction of the effect was similar for the Medium profile but not statistically significant. This result may be attributable to the size of the Low-Medium profile: As the largest profile in number of student members, the profile has more statistical power, allowing for the effect to be significant. However, again it is important not to interpret these results to be causal. Another interpretation might be that higher-performing students tend to be clustered in classes that are not implementing as many student-centered practices, such as Advanced Placement classes, in which students prepare for the exam over the course of the year.

It is also worth remembering that the schools are fairly early in the implementation of PBE. It may be that there is a period during which, as students are exposed to more PBE practices, their standardized assessment scores might be lower. This may be a result of the early phase of implementation and

something that, as the implementation becomes more robust, might yield better scores. We simply do not know, and the matter warrants further, longitudinal research. Regardless of the reason for this difference, if one of the goals of student-centered, proficiency-based education is to positively influence student achievement, then these results might suggest the need to consider appropriate metrics for evaluating the relationship between student-centered, proficiency-based practices and student academic outcomes.

When the interview respondents were asked for whom PBE is most important or who is best served by PBE, we heard a range of responses. Several respondents thought that PBE was particularly important for students who need a little extra time to master content, while they simultaneously questioned PBE's effect on these students' motivation. They also pondered the effect of PBE on more academically motivated and successful students and wondered how the schools could meet these students' needs while addressing the needs of students who needed more time or support. In short, there are many questions about how to implement PBE so that all students benefit.

What do the findings suggest about how to approach the initial implementation of PBE reforms?

Based on the decrease in variability of exposure to different practices that we observe in the Medium profile, it seems that as students experience more exposure to PBE practices, the variation in exposure tends to decrease. This may suggest that as certain PBE-related practices are implemented more widely, overall PBE exposure increases. Although this study does not speculate about what types of practices to initiate first to successfully realize PBE, the results do raise important questions about what types of practices a school or district might choose to implement first versus what types of practices to initiate later in a reform process.

Similarly, the study findings might suggest areas on which to focus professional development and support for teachers. If teachers using most of their class time to deliver lectures is an entrenched practice, what would happen if professional development focused on how to move teachers to employ more diverse strategies beyond lecturing? Would this kind of targeted professional support yield changes in the frequency with which teachers lecture, and might other student-centered, proficiency-based practices, such as group work or project-based learning, increase in turn?

The results of our study suggest that focusing on changing instructional practices is critical to achieving a full transition to student-centered, proficiency-based education and that using the lever of the proficiency-based graduation requirement seems to have yielded much confusion over how to achieve what is arguably the real intent of Maine's current law: that is, to ensure that all students leave high school with the knowledge and skills to be successful. Perhaps if the focus were more squarely on

shifts to instructional practices rather than on changing grading and graduation practices, teachers would have a clearer sense of how they should do their job differently, rather than simply fitting the new grades onto the old model. In addition, if the focus were on shifting instructional practices, it is possible that the community reaction and support for the reform effort would be more positive and would allow for a more authentic shift to the types of practices that are in keeping with the philosophy behind PBE.

Fixsen et al. (2005) describe six stages of implementation: exploration and adoption, program installation, initial implementation, full operation, innovation, and sustainability. In light of the current climate and challenges to the state proficiency-based diploma law, it may be valuable to place the schools in the study within this framework and to remember that the process of implementing a reform as significant as PBE is a long one. Based on our analyses, the schools in the study are in initial implementation of PBE. This is a fragile stage of implementation because everything is so new and implementers are likely to make mistakes as they use newly acquired skills. According to Fixsen et al.,

> *During the initial stage of implementation the compelling forces of fear of change, inertia, and investment in the status quo combine with the inherently difficult and complex work of implementing something new. And, all of this occurs at a time when the program is struggling to begin and when confidence in the decision to adopt the program is being tested. Attempts to implement new practices effectively may end at this point, overwhelmed by the proximal and distal influences on practice and management (e.g., Macallair & Males, 2004).*

The concerns and fears raised by many teachers, students, and parents about how the shift to PBE changes the way learning occurs and school is structured, and about what it means to do well, seem to align with this observation related to early implementation. The hope is that, as is necessary with many reforms, sufficient time will be given to address these fears and to determine a path forward and through the difficult early stages of implementation to realize the intent of the reform.

Appendix A

The following appendix provides information to supplement the analyses for Research Question 1.

Selection of Survey Items for Research Question 1

To select the high-leverage survey items, we relied on theory and the existing literature on proficiency-based learning, as well as on our initial exploratory data analysis. Using current literature as a guide, two team members reviewed all survey items associated with each element (e.g., progression through demonstration of mastery and anytime, anywhere learning) and, for each element, selected three to five items that best represented that element. For example, the literature on mastery-based progression suggests that high-leverage indicators of exposure to opportunities for progression through demonstration of mastery should focus on both the extent to which students are exposed to such instructional practices and the extent to which students are exposed to guidance and feedback from teachers about how mastery will be assessed or determined (Competency Works, 2014; Ryan & Cox, 2017; Steele et al., 2014; Sturgis, 2016). The researchers selected items that reflected these ideas, such as an item that asked students to respond to the prompt "I must show my teachers that I have mastered each standard before I can move on to the next one."

Working separately, and using initial exploratory data analyses as a guide, a third team member used latent profile analysis techniques to explore which items best distinguished among groups of students. The team member ran separate latent profile analyses for five elements (progression through demonstration of mastery, personalization, flexible assessment, development of specific skills and dispositions, and anytime, anywhere learning).[9] The results were then used to identify the final set of 16 high-leverage items, which considers student exposure to implementation of all elements simultaneously.

Details of the Analytic Procedures for Research Question 1

We began process for fitting the LPA models for the 16 high-leverage items by fitting iterative profile models to the data. The goal of this first step of analysis was to arrive at the most likely number of profiles given the data. Once the most plausible profile solution was established, the absolute measures of model fit were examined. These statistics provide an indication of whether the profile solutions are meaningfully distinct from one another and give a sense of the level of certainty with which students were assigned to their most likely PBE exposure profile. We estimated plausible models

[9] Although the Competency-based Learning Survey includes only four distinct domains, based on the extant literature on student-centered, proficiency-based learning, we decided to add the element of anytime, anywhere learning as its own element in the LPA analyses.

under the single-level, cluster robust standard error assumptions using MPlus version 8 (Muthén & Muthén, 2015).[10]

To arrive at the most likely profile solution, we fit successive (k +1) models to the data and compared the fit of these models with the best-fitting previous profile solution. The ideal number of profiles was discerned using fit statistics recommended for use with latent mixture models (Jung & Wickrama, 2008), such as: Bayesian Information Criterion (BIC) and its sample adjusted estimate, as well as the Akaike Information Criterion (AIC)[11], relative entropy value (ranging from zero to one) > .70, indicating classification accuracy; and profile size (comprising at least 1% of the sample to ensure generalizability and replicability) (Nylund, Asparouhov, & Muthen, 2007).

Once the most plausible profile solution was established, the absolute measures of model fit were examined. These statistics provide an indication of whether the profile solutions are meaningfully distinct from one another. Because the LPA is a probabilistic representation of the data, it is critical to understand not just the most likely profile solution but also the degree of certainty with which individuals are assigned to their most likely profile. Four different absolute measures of model fit were examined. First, the model estimated proportions for class assignment (Pr_k) were examined. These are simply the model estimated proportions for class assignment. It can be useful to look at the 95% confidence intervals to examine how tight the bands are around the estimated proportions. These values are also used to calculate other summary statistics. The average posterior class probability ($AvePP_k$) enables evaluation of the classification uncertainty for each of the latent classes separately. It is calculated by assigning individuals to their modal profile probability and calculating the average of the probabilities within profiles. Nagin (2005) suggests that all $AvePP_k$ values greater than .7 indicate adequate separation and classification precision of the profiles. The odds of correct classification ratio (OCC_k) is provided to give a sense of the assignment accuracy of the model solution. OCC_k values = 1 for a class indicate that the probability of assignment for members of that class are no better than chance. Nagin (2005) suggests that an $OCC_k > 5.0$ for all groups indicates that the model has good separation between latent classes and high accuracy in the assignment of individuals to latent classes. Finally, the modal class assignment proportion ($mcaP_k$) is the proportion of individuals in the sample who are modally assigned to a particular profile based on their posterior probability of profile assignment. If individuals were assigned to a profile with perfect certainty, then $mcaP_k$ would be equal to the model estimated Pr_k. Larger discrepancies between the two statistics are indicative of larger, latent class assignment errors. As such, we look for alignment between this statistic and the model estimated Pr_k.

[10] Bias corrections to the standard errors account for the fact that students were nested in schools.

[11] Models with lower fit statistics are considered more probable given the data.

Supplemental Results for Research Question 1

We fit a series of profile models ranging from 1 to 4 implementation profiles. The estimated LPA models were compared using fit statistics recommended for use with latent mixture models, which latent profile analysis models. These fit statistics are provided in Table A1. The final model, the 3-profile solution, showed an entropy of .80. It should be noted that although the 4-profile model showed lower fit statistics, subsequent exploration of the model solution indicated that some of the profile solutions exhibited mixing, which led to some lower probability of profile assignment, lower entropy levels and small class sizes. As such, we favored the stability of the 3-profile solution.

Table A1. Model fit statistics for the k+1 sequence of model tests

Model	No. of free parameters	BIC	Sample Adjusted BIC	AIC
1-profile	32	73832.435	73730.773	73656.172
2-profile	49	69590.993	69435.322	69321.089
3-profile	66	68519.624	68309.945	68156.080
4-profile	83	68250.705	67987.018	67793.521

Note. All models use cluster robust standards errors. BIC = Bayesian information criterion. AIC = Akaike information criterion.

Technical Quality of the 3-Profile Solution for Research Question 1

To investigate the technical quality of the 3-profile solution, we examined the absolute measures of model fit. LPA is based on an explicit model of the data, which reflects the fact that recovered latent profiles are uncertain; a student's membership in any given latent profile is reflected in a probability that sums to 1 across the total number of profiles modeled. As such, a student has an assigned probability of membership in each of the estimated latent profiles to various degrees of certainty. Profile solutions that exhibit adequate technical quality assign students to their respective modal profile with certainty. The absolute measures of model fit are indicators of the quality of the 3-profile solution and are provided in Table A2. Overall, these values indicate that the 3-profile solution provide adequate separation of the latent profiles and that students were assigned to their modal profile with a high degree of certainty. The model estimated marginal proportions (Pr_k) for class assignment indicated that most students (51%) were assigned to the Low implementation PBE profile, followed by the Medium implementation PBE profile (31%) and the High implementation PBE profile (19%). The modal class assignment proportion ($mcaP_k$) values were nearly identical to the model estimated marginal proportions, which indicates that students were assigned to their PBE implementation profiles with a high degree of certainty. In addition, all $AvePP_k$ values are above .7, which indicates that there was adequate separation between the three profiles. Finally, all OCC_k values are above 5 for all latent profiles, which further supports the notion that the assignment of students to their most likely implementation profiles contains little error, and that there is good separation between the latent classes. In sum, three implementation profiles represent distinct subgroups of students who report experiencing the implementation of proficiency-based education (PBE) similarly within profiles and differently between the profiles.

Table A2. Absolute measures of model fit for the 3-profile solution

Profile	Pr_k	Lower 95%	Upper 95%	$mcaP_k$	$AvePP_k$	Variance $AvePP_k$	Standard Deviation $AvePP_k$	OCC_k
Minimal PBE	0.31	0.29	0.33	0.31	0.91	0.02	0.14	23.76
Low-Medium PBE	0.51	0.48	0.53	0.51	0.91	0.02	0.13	9.30
Medium PBE	0.19	0.17	0.21	0.18	0.91	0.02	0.14	40.95

Supplemental Information for Survey Items for Research Question 1

The following table provides information useful for interpreting figure 2. Specifically, Table A3 provides information about the survey questions, item group (that is subscale) and rating scale used for each item in the survey.

Table A3. Description of survey items and rating scale

Item Name	Item Description	Item Group	Rating Scale
ONLINECOMP	I am able to complete some or all of the course requirements online.	*Anytime, Anywhere Learning*	0 = No courses 1 = Some courses 2 = All or most courses
PROJCRED	If I complete a project that wasn't assigned at school but is related to a course I am taking, I can earn credit for the project in that course.	*Anytime, Anywhere Learning*	0 = No courses 1 = Some courses 2 = All or most courses
INTERNCRED	I can earn credit for completing an internship or job-shadowing in the community.	*Anytime, Anywhere Learning*	0 = No courses 1 = Some courses 2 = All or most courses
GIVLECTUR	My teachers spend most of class time giving a lecture or presentation to the whole class.	*Personalization*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
SMALLGRP	My teachers work with students in small groups or individually.	*Personalization*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
TEACHDIFF	My teachers teach the material in several different	*Personalization*	0 = Never 1 = Seldom

Item Name	Item Description	Item Group	Rating Scale
	ways in order to help students learn.		2 = Sometimes 3 = Often 4 = Always
CHOOSEHOW	I have had opportunities to choose how to show my teachers what I have learned.	*Flexible Assessment*	0 = Not at all 1 = 1-2 times 2 = 3-4 times 3 = 5 or more times
RETAKE	If I do poorly on an assignment on the first try, I can try again later.	*Flexible Assessment*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
DEMOLEARN	To show that I have mastered a course standard, I must demonstrate my learning in more than one way.	*Flexible Assessment*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
ADVICE	When I have trouble learning something new, my teachers give me advice and strategies that help me to stick with it.	*Ownership/Agency*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
TRACKPROG	Teachers show students how to keep track of their progress on each of the standards.	*Ownership/Agency*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
ONTIME	Teachers show students strategies for making sure all assignments are completed on time.	*Ownership/Agency*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
PREPGRAD	I know which steps to take during high school in order to prepare for what to do after I graduate.	*Ownership/Agency*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
MASTERCOMP	I must show my teachers that I have mastered each standard before I can move on to the next one.	*Progression Through Demonstration of Mastery*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always
NEXTCOMP	I am able to move on to the next standard when I am ready, even if other students in the course are not ready.	*Progression Through Demonstration of Mastery*	0 = Never 1 = Seldom 2 = Sometimes 3 = Often

Item Name	Item Description	Item Group	Rating Scale
RUBRIC	My teachers give me a rubric so that I know how I am progressing on each standard.	*Progression Through Demonstration of Mastery*	4 = Always 0 = Never 1 = Seldom 2 = Sometimes 3 = Often 4 = Always

Information about the Analytic Approach used for Research Question 1B

The "3-step" approach was used to estimate multinomial models (Asparouhov & Muthén, 2014).[12] Models were generated using full-information maximum likelihood estimation, and a sandwich estimator was used to generate conservative standard errors that accounted for students clustered in schools. All of the student characteristic variables were entered into the model in a single block. Significance of predictors was determined with a Wald statistic.

[12] The first step in the 3-step approach was to estimate the latent profile model. Next, a nominal most likely profile variable was created. Next, a mixture mode that used the most likely profile variable, the profiles, and the predictors was estimated. Regression estimates were then generated that accounted for the quality of classifications of students into their respective PBE implementation profiles. The 'r3step' function in MPlus version 8 was used to conduct the aforementioned steps automatically.

Appendix B

The following appendix lists information about the 12 teachers who were interviewed for this study.

Table B1. Teachers Interviewed

School	Role	Years of experience	Rating of PBE on 1-10 scale
Princeton	Special Education	6	2
Princeton	English	4	9
Princeton	Social Studies	3	8
Princeton	Science	5	7
Princeton	Math	28	"Somewhere in the middle"
Walden	Social Studies	34	8
Walden	Math	34	2
Walden	English	7	5
Fields	Social Studies	26	4
Fields	Math	8	8-9
Fields	English	26	8
Fields	Science	16	10

APPENDIX C

Competency-based Learning Survey with Engagement Scale

The following survey was distributed to students in grades 9–12 who attended 11 secondary schools located within 10 districts in rural Maine during the 2016/17 academic school year. This survey is published on the IES website[13] in its entirely, excluding the engagement items, which were added to the survey for this study. Language in the survey below was adapted from the original survey to meet the specific context of the schools in Maine that were the subject of this study.

This is a survey about your experiences at this school, including courses at this school. All of your answers will be confidential. This survey is NOT A TEST and there are no right or wrong answers. We really hope you will try to answer every question, but you may skip any question you do not wish to answer. It should take about 20 minutes to complete the survey. By clicking on **"NEXT"** you agree to participate in this study.

Module A: Demographics

Are you male or female?

- ○ Male
- ○ Female
- ○ Prefer not to say

---

[13] Ryan, S., & Cox, J. (2016). *Guide to the Competency-based Learning Survey for Students* (REL 2016–165). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

What is your race or ethnicity?

- ○ White
- ○ Black or African American
- ○ American Indian or Alaska Native
- ○ Asian
- ○ Native Hawaiian or Pacific Islander
- ○ Other

How much education do you expect to complete?

- ○ Less than high school
- ○ High school diploma
- ○ GED
- ○ Associate's degree
- ○ Some college but less than a bachelor's degree
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctoral or advanced professional degree

Module B: Student Understanding and Beliefs

This set of questions asks about practices and requirements at your school. Your responses to these questions will help school leaders and teachers understand students' experiences at your school. As is true of all questions in this survey, there are no right or wrong answers to these questions.

For the following statements, indicate how much you agree or disagree.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Not sure
The graduation requirements at my school will prepare me for what I want to do after high school	○	○	○	○	○
Students should get more than one opportunity to show whether they have learned the important course material	○	○	○	○	○
Students should get more than one opportunity to pass a test or exam	○	○	○	○	○
Homework is important to complete even if it is not GRADED	○	○	○	○	○
If two students in the same course complete different types of assignments, they should still have the opportunity to earn the same grade in the course	○	○	○	○	○
My grades are a good reflection of what I have learned	○	○	○	○	○
Most colleges will know what my course grades mean. In other words, most colleges will know how to evaluate the grades I've received in the courses I've taken at my school	○	○	○	○	○

For some of the next several items, several options may seem **SOMEWHAT TRUE**. However, you should choose the response that is the **MOST TRUE** at your school.

In courses at my school, students must show their learning on each standard in more than one way. For example, students must show that they have met or are proficient on the standard on more than one assignment, assessment, or exam.

- ○ Often
- ○ Sometimes
- ○ Never
- ○ Don't know

Why do you think students must show their learning on each standard in more than one way? Choose the response that seems **MOST TRUE** about this school.

- ○ A student's grade in the course should be based on multiple scores
- ○ Teachers need multiple opportunities to see whether students have learned the material
- ○ Students stay on task better if they have enough work to do
- ○ Completing multiple assignments and tests helps prepare students for the state standardized test

In courses at my school, students are able to choose how they want to show what they have learned from several different options. For example, options such as taking a test, writing a paper, and completing a project, etc.

- ○ Often
- ○ Sometimes
- ○ Never
- ○ Don't know

Why do you think students in courses at this school are able to choose how they want to show their learning? Choose the response that seems **MOST TRUE** about this school.

- ○ Most students know how they learn best
- ○ Some students aren't good at taking tests/exams
- ○ Each student should be able to get an A in the course
- ○ It is good for students to have variety in their work

Students at my school are able to progress at their own individual pace in courses.

- ○ Often
- ○ Sometimes
- ○ Never
- ○ Don't know

Why do you think students in courses at this school are able to progress at their own individual pace? Choose the response that seems **MOST TRUE** about this school.

- ○ Some students fall behind if they have been absent a lot
- ○ Some students are less interested in certain topics
- ○ Some students may need different amounts of time to learn the material
- ○ Some students don't complete all of their work on time

Do any of the courses at your school use a proficiency-based grading system? In other words, at the end of the course, do you get a proficiency-based score or grade instead of (or along with) a traditional letter grade?

- ○ Yes
- ○ No
- ○ Don't know

Why do you think all or some of the courses at your school use a proficiency-based grading system? Choose the response that is **MOST TRUE** at your school.

- ○ Teachers do a better job of grading student work under a proficiency-based grading system
- ○ Students take fewer tests under a proficiency-based grading system
- ○ A proficiency-based grading system makes it easier for all students to graduate from high school
- ○ A proficiency-based grading system provides better information about what a student has learned

---

Students at my school are assigned homework.

- ○ Often
- ○ Sometimes
- ○ Never
- ○ Don't know

Why do you think students at this school are assigned homework? Choose the response that seems MOST TRUE at this school.

- ○ Homework provides students with opportunities to practice a skill before being assessed on that skill
- ○ Students in high school should have homework almost every night in order to learn the material
- ○ Teachers don't have time to teach all of the important material during class
- ○ Completing several hours of homework most nights helps students get ready for college

---

At this school, when teachers talk about standards, they are referring to:

- ○ The courses a student must take to graduate
- ○ The important skills and knowledge a student must learn to graduate
- ○ The required credits a student must complete to graduate
- ○ I'm not sure what teachers mean when they talk about standards
- ○ Teachers do not talk about standards

Do you have to demonstrate proficiency on specific standards in any of your courses in order to graduate from your high school?

- ○ Yes
- ○ No
- ○ Don't know

Has a principal or teacher explained to you why it is important to master specific standards in order to graduate from your high school?

- ○ Yes
- ○ No
- ○ Don't know

Module 2: Progression through demonstration of mastery

In this section of the survey, you will answer some questions related to courses at this school, including how students make progress in their learning. Again, there are no right or wrong answers to these questions.

For these items, indicate about how often this happens in your courses at this school.

	Always	Often	Sometimes	Seldom	Never	Not sure
I know what I need to do to show my teacher that I am making progress on each standard	○	○	○	○	○	○
I must show my teachers that I have met or am proficient on a standard before I can move on to the next standard	○	○	○	○	○	○
I am able to move on to the next standard when I am ready, even if other students in the course are not ready	○	○	○	○	○	○
Students in my courses work on the same standard at the same time	○	○	○	○	○	○

For these items, indicate about how often this happens in your courses at this school.

	Always	Often	Sometimes	Seldom	Never	Not sure
I understand how the material I am studying in my courses will help me in the future	○	○	○	○	○	○
My teachers share examples of excellent student work	○	○	○	○	○	○
My teachers let me know how my work will be assessed or graded for each standard	○	○	○	○	○	○
My teachers give me a rubric so that I know how I am progressing on each standard	○	○	○	○	○	○

Module C: Personalization

Now you will answer some questions about different options students have for earning credit in courses at this school. You will also answer some questions about different ways in which teachers work with students. Again, there are no right or wrong answers to these questions.

Think about the courses you are taking right now. For about how many of your courses is each of the following options true?

	All or most	Some	None	Not sure
I am able to complete some or all of the course requirements online	○	○	○	○
If I complete a project that wasn't assigned at school but is related to a course I am taking, I can earn credit for the project in that course	○	○	○	○
I can earn credit for taking courses at another high school	○	○	○	○
I can earn credit for taking courses at a college (for example, "dual-credit courses")	○	○	○	○
I can earn credit for completing an internship or job-shadowing in the community	○	○	○	○

For these items, indicate about how often this happens in courses at this school.

	Always	Often	Sometimes	Seldom	Never	Not sure
Students in my courses all work on the same assignment at the same time	○	○	○	○	○	○
My teachers spend most of class time giving a lecture or presentation to the whole class	○	○	○	○	○	○
My teachers work with students in small groups or individually	○	○	○	○	○	○
My teachers notice if I need extra help	○	○	○	○	○	○
My teachers teach the material in several different ways in order to help students learn	○	○	○	○	○	○

For these questions, think about how often each event occurred DURING THE CURRENT SEMESTER.

	5 or more times	3-4 times	1-2 times	Not at all	Not sure
My teachers or a counselor/advisor discussed how I am doing on each standard	○	○	○	○	○
My teachers gave me written feedback on my work	○	○	○	○	○
I have had opportunities to choose how to show my teachers what I have learned	○	○	○	○	○

Module D: Flexible Assessment

Thanks for being honest in your answers to the survey questions. Remember, there are no right or wrong answers. In this section of the survey, you will answer some questions about how you are asked to show your learning in your courses. In other words, these questions ask about different kinds of assessments used at this school.

For these questions, think about how often each event occurred DURING THE CURRENT SEMESTER.

	5 or more times	3-4 times	1-2 times	Not at all	Not sure
I have created drawings or models to show what I have learned	○	○	○	○	○
I have taken tests or quizzes to show what I have learned	○	○	○	○	○
I have given a performance to show what I have learned (for example, performing in a video or skit/play, playing an instrument)	○	○	○	○	○
I have given a presentation to show what I have learned	○	○	○	○	○
I have completed a project at school to show what I have learned	○	○	○	○	○
I have completed a project in the community to show what I have learned	○	○	○	○	○

For each item, indicate about how often this happens in the courses you are taking.

	Always	Often	Sometimes	Seldom	Never	Not sure
If I do poorly on an assessment on the first try, I can try again later	○	○	○	○	○	○
To show that I have met or that I'm proficient in a course standard, I must demonstrate my learning in more than one way	○	○	○	○	○	○

Module E: Skills and Dispositions

Thanks for your responses. You're almost done!

For each item, indicate about how often this happens in courses you are taking.

	Always	Often	Sometimes	Seldom	Never	Not sure
Teachers show or explain strategies students can use to keep track of their progress on each of the standards	○	○	○	○	○	○
Teachers show or explain strategies students can use to complete all assignments and assessments on time	○	○	○	○	○	○
Teachers encourage students to take responsibility for their work	○	○	○	○	○	○
I know which steps to take during high school in order to prepare for what I want to do after I graduate	○	○	○	○	○	○
If I need information that I don't have in order to complete an assignment, I know where to get it	○	○	○	○	○	○

If it is difficult for me to get an assignment done on my own, I know strategies I can use so that I'm sure to get the work finished on time	○	○	○	○	○	○
Teachers show or explain strategies students can use to help each other learn	○	○	○	○	○	○
Teachers show or explain strategies students can use to work together successfully in groups	○	○	○	○	○	○
Teachers encourage students to help each other outside of class	○	○	○	○	○	○

This section includes a few more questions about teaching and learning at this school. Again, there are no right or wrong answers to these questions.

For each item, indicate about how often this happens in the courses you are taking.

	Always	Often	Sometimes	Seldom	Never	Not sure
Teachers encourage students to respect the feelings of others	○	○	○	○	○	○
Teachers show or explain to students how to treat each other with respect	○	○	○	○	○	○
Teachers explain to students how they can disagree with each other respectfully	○	○	○	○	○	○
When I have trouble learning something new, my teachers give me advice and strategies that help me to keep trying	○	○	○	○	○	○
My teachers notice when I take extra time and effort on something that is difficult for me	○	○	○	○	○	○
If I get a low score on an assessment, my teachers help me figure out how I can still do well in the class	○	○	○	○	○	○

Engagement

This page and the one that follows it include the last set of questions.

For the following statements, indicate how much you agree or disagree.

	Strongly disagree	Disagree	Agree	Strongly Agree
I feel proud of being part of my school.	○	○	○	○
I am treated with as much respect as other students in my class.	○	○	○	○
I can get a good job even if my grades are bad.	○	○	○	○
The only time I get attention in school is when I cause trouble.	○	○	○	○
I like to participate in a lot of school activities (for example, sports, clubs, and plays).	○	○	○	○
School is one of the most important things in my life.	○	○	○	○
Many of the things we learn in class are useless.	○	○	○	○
Most of my teachers don't really care about me.	○	○	○	○
Most of the time I would like to be any place other than in school.	○	○	○	○
There are teachers or other adults in my school that I can talk to if I have a problem.	○	○	○	○
Most of what I learn in school will be useful when I get a job.	○	○	○	○

School is one of my favorite places to be.	○	○	○	○
People at school are interested in what I have to say.	○	○	○	○
School is often a waste of time.	○	○	○	○
Dropping out of school would be a huge mistake for me.	○	○	○	○
School is more important than most people think.	○	○	○	○

Appendix D

The following appendix provides information to supplement the analyses for Research Question 2.

Sequence of Model Testing for Engagement and SAT Models

Students were designated to a specific PBE exposure profile (Minimal, Medium-Low, Medium) using a "classify then analyze" approach, where students were assigned to the profile for which they had the highest probability of being a member. This approach performs optimally when modal class probabilities are high, as was evidenced in research question 1A (Dziak, Bray, Zhang, Zhang, & Lanza, 2016).

To determine the extent to which membership in a specific profile is related to academic and engagement outcomes, while accounting for other factors (such as school membership, IEP status, FRL status, etc.), we fit separate models for students' self-reported engagement outcomes and for their performance on the SAT with varying intercepts specified to account for potential variation in the outcomes between schools. The sample used in the SAT model was constrained to 11th graders in the 2016/17 school year. As described above, we included only 11th graders in the outcome analyses because we wanted to include only those students who would have taken the SATs in the same year as the survey was administered and not include 12th graders, for whom the SAT could not be considered an outcome measure because they might have taken the SAT in the prior year. Both students' engagement and their SAT scores were converted to z-scores prior to estimating the models. Continuous inputs were grand-mean centered and scaled by two times their standard deviations, such that estimates generated are then directly comparable for untransformed binary predictors. The analyses of students' engagement outcomes used all students in the sample. Models were estimated using full-information maximum likelihood estimation.

Model testing for both outcomes followed the "bottom-up" approach proffered by Hox, Moerbeek, and van de Schoot (2010), using a 2-level multilevel model, with students nested in schools. The model testing commenced in sequence, where predictors were entered into the model in blocks. The goal of this approach was to isolate the variation in the outcomes possibly attributed to the binary indicators for students' membership in PBE exposure profiles while accounting for variance in the outcomes associated with student characteristics. Overall, the sequence of model testing was very similar for both the engagement and the SAT models, so these are described jointly, with minor differences between the two models noted. For both models, the first sequence of model testing was to establish whether there were potential school effects. This was achieved by estimating the intraclass correlation (ICC). The ICC can be thought of as the correlation of scores on the outcome measure among students

within the same school. Next, student characteristics variables were staged into both models as follows:

- **Sequence of model testing for the analysis of student engagement data**
  - Model 1: gender, race, grade level, free and reduced-price lunch status, IEP status, and the number of unexcused absences
  - Model 2: unweighted GPA
  - Model 3: PBE exposure profiles
- **Sequence of model testing for the analysis of student SAT data**
  - Model 1: gender, race, grade level, free and reduced-price lunch status, IEP status, and the number of unexcused absences, unweighted GPA
  - Model 2: PSAT scores
  - Model 3: PBE exposure profiles

The best fitting models were determined using the BIC statistic,[14] with models containing a lower BIC value taken as evidence for a more plausible model given the data (Raftery, 1995). Once the best-fitting model was established, the model residuals were examined for opportunities to improve model fit through the inclusion of polynomial terms for continuous predictors (i.e., unexcused absences, unweighted GPA and PSAT scores). In the final models, two coefficients were examined to answer Research Question 2 directly: the binary indicators for whether students were in the Low-Medium or Medium exposure profiles. These estimates represent the association between students' PBE exposure and the outcomes while accounting for variance in the outcomes associated with other student characteristics and all other variables in the model.

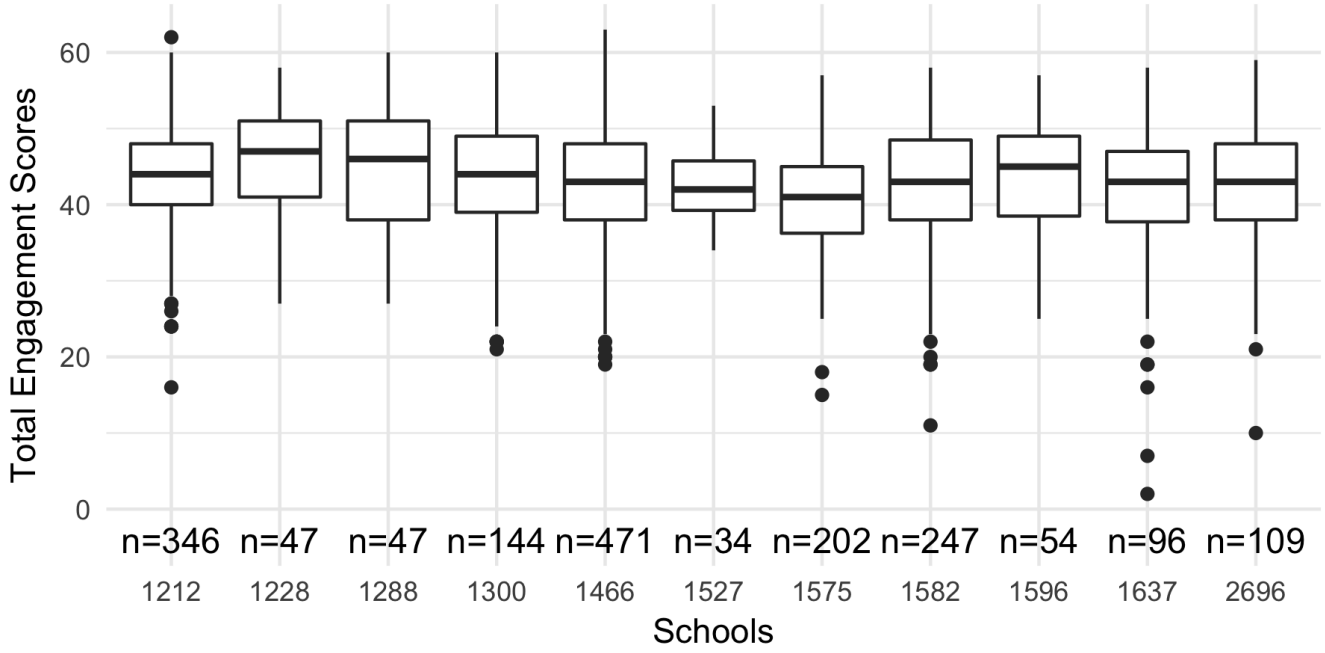
Supplemental Results for the Analysis of Engagement Scores

Variation in students' engagement scores

A plot of students' engagement scores by school is provided in Figure D1; this provides context to support the ICC presented for Research Question 2. The plot shows very little variation in students' engagement scores between schools, which supports the finding that being a member of a particular school is not associated with students' engagement scores.

[14] Loglikelihood ratio tests were not used because of differences that arise in the sample size of models used for the various combinations of model predictors. These small differences in sample size mean that models cannot be considered truly nested and that the loglikelihood ratio test is not valid for use. The BIC statistic is a suitable approach to establishing plausible models given the data (Raftery, 1995),

Figure D1. Variation in Students' Engagement Scores by School



Extending the regression models for students' engagement scores

After establishing the variance in the outcome that could be attributed to schools (i.e., the ICC statistic), we commenced with the sequence of model testing described in the Methodology section. Our interest in the sequence of model testing was to observe changes in the pseudo R^2 values and BIC statistics that led up to the model that ultimately included the indicator variables for students' membership in PBE implementation profiles. The goal of this sequence of model testing was to establish whether, after student characteristic information was included in the model, the indicators for PBE profiles were predictive of students' engagement scores. The smaller BIC values for models 3 and 4 in Table D1 provided support for the contribution of PBE profiles in predicting students' engagement scores.

Table D1. Pseudo R^2 values and BIC statistics for regression models of students' engagement scores

Model	Pseudo R^2 value	BIC
Model 0: Null model	N/A	2618.093
Model 1: Student characteristics (i.e., gender, race, grade level, free and reduced-price lunch status, IEP status)	6.8%	2120.11
Model 2: Student characteristics + Unweighted GPA	12.3%	1875.826

Model 3: Student characteristics + Unweighted GPA + PBE Profiles	25.1%	1682.698
Model 4: Student characteristics + Unweighted GPA + PBE Profiles + Unweighted GPA^2	25.6%	**1681.859**

Table D2 provides the complete set of regression estimates, standard errors, and variance components for each sequence of model testing.

Table D2. Results from the regression model for students' engagement scores

	B	CI	p	B	CI	p	B	CI	p	B	CI	p	B	CI	p
Fixed Parts															
(Intercept)	0.11	0.02 – 0.20	0.018	0.07	−0.04 – 0.17	0.229	−0.18	−0.30 – −0.06	0.004	−0.22	−0.34 – −0.09	<.001	−0.23	−0.36 – −0.10	<.001
Male	−0.06	−0.11 – −0.01	0.025	−0.01	−0.06 – 0.04	0.681	−0.03	−0.08 – 0.02	0.21	−0.03	−0.07 – 0.02	0.305	−0.03	−0.07 – 0.02	0.303
Gender Not Reported	−0.23	−0.44 – −0.03	0.028	−0.21	−0.41 – −0.00	0.05	−0.21	−0.40 – −0.02	0.03	−0.2	−0.39 – −0.01	0.037	−0.2	−0.39 – −0.01	0.039
Student Is Not White	−0.06	−0.14 – 0.03	0.169	−0.06	−0.14 – 0.03	0.184	−0.05	−0.13 – 0.03	0.201	−0.05	−0.13 – 0.03	0.215	−0.05	−0.12 – 0.03	0.25
10th Grade vs. 9th Grade	−0.06	−0.13 – 0.00	0.065	−0.08	−0.15 – −0.01	0.035	−0.06	−0.12 – 0.01	0.1	−0.04	−0.11 – 0.02	0.206	−0.04	−0.11 – 0.02	0.198
11th Grade vs. 9th Grade	−0.08	−0.15 – −0.01	0.018	−0.08	−0.15 – −0.01	0.027	−0.01	−0.08 – 0.06	0.746	0	−0.06 – 0.07	0.956	0	−0.07 – 0.07	0.989
12th Grade vs. 9th Grade	−0.01	−0.08 – 0.07	0.876	−0.02	−0.09 – 0.05	0.595	0.02	−0.05 – 0.09	0.606	0.03	−0.04 – 0.10	0.377	0.03	−0.04 – 0.10	0.368
Student receives FRL	−0.09	−0.15 – −0.04	<.001	−0.04	−0.09 – 0.02	0.191	−0.06	−0.11 – −0.01	0.024	−0.05	−0.10 – −0.00	0.049	−0.05	−0.10 – 0.00	0.057
Student has IEP	−0.02	−0.09 – 0.05	0.547	0.02	−0.06 – 0.09	0.668	−0.08	−0.15 – −0.01	0.02	−0.08	−0.15 – −0.01	0.029	−0.08	−0.14 – −0.01	0.033
Unexcused days absent	−0.13	−0.19 – −0.08	<.001	−0.04	−0.10 – 0.02	0.191	−0.04	−0.09 – 0.02	0.221	−0.04	−0.10 – 0.01	0.137	−0.13	−0.24 – −0.03	0.012
Unweighted GPA				0.27	0.21 – 0.33	<.001	0.29	0.23 – 0.34	<.001	0.33	0.27 – 0.39	<.001	0.32	0.25 – 0.38	<.001
Low-Medium PBE Profile vs. Minimal PBE Profile							0.31	0.26 – 0.36	<.001	0.31	0.26 – 0.37	<.001	0.31	0.26 – 0.36	<.001
Medium PBE Profile vs. Minimal PBE Profile							0.52	0.45 – 0.59	<.001	0.52	0.45 – 0.59	<.001	0.51	0.44 – 0.59	<.001
Unweighted GPA^2										0.07	0.02 – 0.11	0.005	0.06	0.02 – 0.11	0.005
Unexcused Days Absent^2													0.05	0.00 – 0.10	0.041
Random Parts															
σ^2 (L1 Residual)		0.23			0.218			0.186			0.185			0.184	
$\tau_{00,\text{School\_ID}}$ (L2 Random Intercept)		0.008			0.013			0.02			0.02			0.02	
$N_{\text{School\_ID}}$		9			9			9			9			9	
$ICC_{\text{School\_ID}}$		0.034			0.058			0.098			0.1			0.098	
Observations		1474			1344			1344			1344			1344	
R^2 / Ω_0^2		.068 / .068			.123 / .123			.251 / .251			.256 / .255			.258 / .258	

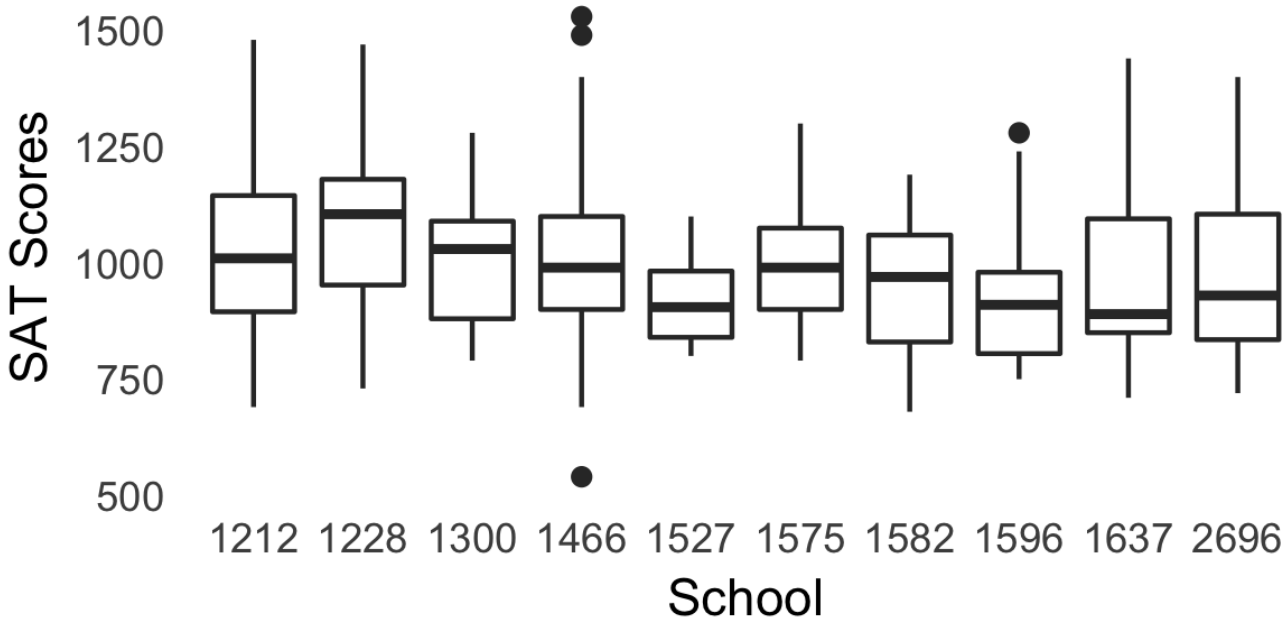
Note. Continuous inputs have been transformed by subtracting the mean and dividing by 2 standard deviations.

Supplemental Results for the Analysis of SAT Scores

Variation in students' SAT scores

A plot of students' SAT scores by school is provided in Figure D2; this provides context to support the ICC presented for Research Question 2. The plot shows very little variation in students' SAT scores between schools, which supports the finding that for students, being a member of a particular school is not associated with engagement scores.

Figure D2. Variation in 11th Grade Students' SAT Scores by School



Extending the regression models for students' SAT scores

After establishing the variance in the outcome that could be attributed to schools (i.e., the ICC statistic), we commenced with the sequence of model testing described in the Methodology section. Model fit statistics are provided in Table D3. Two results are notable. The first is that the inclusion of student characteristic data and PSAT scores in the model produced a model (i.e., Model 2) that accounted for 81.1% of the variance in SAT scores. The inclusion of the implementation profiles as a predictor of students' SAT scores in the model (i.e., Model 3) did not produce a BIC statistic that was smaller than that of the model containing just student characteristic data and SAT scores. However,

residuals from this model suggested that the inclusion of a quadratic term for students' PSAT scores was appropriate.

Table D3. Pseudo R^2 values and BIC statistics for regression models of students' SAT scores

Model	Pseudo R^2 value	BIC
Model 0: Null model	N/A	589.304
Model 1: Student characteristics (i.e., gender, race, grade level, free and reduced-price lunch status, IEP status)	77.1%	25.51128
Model 2: Student characteristics + PSAT	81.1%	17.95774
Model 3: Student characteristics + PSAT + PBE Profiles	81.5%	19.08465
Model 4: Student characteristics + PSAT + PBE Profiles + $PSAT^2$	82%	**15.84227**

Table D4 provides the complete set of regression estimates, standard errors, and variance components for each sequence of model testing.

Table D4. Results from the regression model for students' SAT scores

	B	CI	p	B	CI	p	B	CI	p	B	CI	p	B	CI	p	B	CI	p
Fixed Parts																		
(Intercept)	-0.01	-0.07 - 0.06	0.881	0.06	0.02 - 0.09	0.004	0.03	-0.04 - 0.09	0.371	0.09	0.02 - 0.15	0.012	0.07	0.00 - 0.13	0.041	0.07	-0.00 - 0.14	0.068
PSAT				0.88	0.83 - 0.94	<.001	0.72	0.65 - 0.80	<.001	0.71	0.64 - 0.79	<.001	0.67	0.59 - 0.75	<.001	0.67	0.59 - 0.75	<.001
Male							0.06	0.00 - 0.12	0.035	0.06	0.00 - 0.12	0.048	0.05	-0.01 - 0.11	0.081	0.05	-0.01 - 0.11	0.081
Gender Not Reported							0.45	0.14 - 0.76	0.005	0.42	0.11 - 0.73	0.009	0.37	0.06 - 0.68	0.021	0.37	0.06 - 0.68	0.021
Student Not White							-0.03	-0.12 - 0.07	0.584	-0.04	-0.13 - 0.06	0.462	-0.04	-0.14 - 0.05	0.363	-0.04	-0.14 - 0.05	0.369
Student Receives FRL							-0.02	-0.08 - 0.05	0.596	-0.02	-0.09 - 0.04	0.495	-0.02	-0.08 - 0.04	0.497	-0.02	-0.08 - 0.04	0.512
Student Has IEP							-0.24	-0.34 - -0.13	<.001	-0.22	-0.32 - -0.13	<.001	-0.26	-0.36 - -0.16	<.001	-0.26	-0.36 - -0.16	<.001
Unexcused Absences							-0.06	-0.17 - 0.04	0.253	-0.06	-0.16 - 0.04	0.265	-0.04	-0.14 - 0.05	0.381	-0.05	-0.22 - 0.12	0.551
Unweighted GPA							0.19	0.10 - 0.29	<.001	0.19	0.10 - 0.28	<.001	0.19	0.10 - 0.28	<.001	0.19	0.10 - 0.28	<.001
Low-Medium PBE vs. Minimal PBE										-0.1	-0.16 - -0.04	0.002	-0.1	-0.16 - -0.04	<.001	-0.1	-0.16 - -0.04	<.001
Medium PBE vs. Minimal PBE										-0.08	-0.18 - 0.02	0.129	-0.09	-0.19 - 0.01	0.078	-0.09	-0.19 - 0.01	0.078
$PSAT^2$													0.13	0.04 - 0.21	0.003	0.13	0.04 - 0.21	0.003
$Unexcused\ Absences^2$																0.01	-0.15 - 0.17	0.91
Random Parts																		
σ^2		0.246			0.058			0.048			0.046			0.045			0.045	
$\tau_{00,School\_ID}$		0.004			0.001			0.003			0.001			0.001			0.001	
N_{School_ID}		10			10			8			8			8			8	
ICC_{School_ID}		0.014			0.023			0.051			0.031			0.02			0.02	
Observations		395			294			240			240			240			240	
R^2 / Ω_0^2		.034 / .021			.771 / .771			.809 / .809			.815 / .815			.820 / .820			.820 / .820	

Note. Results are constrained to 11th graders. Continuous inputs have been transformed by subtracting the mean and dividing by 2 standard deviations.

Technical details for the regression models examining moderation for Research Question 2A

Fitting a multilevel varying intercept model allowed for the slope of the estimate for implementation profiles to vary by student characteristics. Positive interaction terms were indicative of a relationship between the specific PBE profiles that was stronger for the student characteristic variable specified in the interaction, whereas a negative interaction indicated that the association between a PBE exposure profile and the outcomes was weaker for the student characteristic variable specified in the interaction. The significance of any interaction terms was taken as evidence for a moderated effect for the PBE exposure profiles. Models were fit separately for the engagement and SAT outcomes. These models built on the regression models for Research Question 2 by specifying interactions between the PBE exposure profiles and the following variables: gender, race, grade level, FRL status, and IEP status. Interactions were staged in their respective regression models individually. Overall model fit for each model compared with the model without the interaction was compared using the BIC statistic and Bayes Factor.

APPENDIX E

The following appendix includes the protocols used for focus groups with parents and students and for interviews with teachers, superintendents, and principals. Note that our interview protocol format changed after we conducted the superintendent interviews, to create greater capacity to examine common information across sites.

Interview and focus group protocols

Superintendent Interview Protocol

Context:
Tell me about the community or communities that your high school serves.
Probes:

How would you describe the educational attainment level among adults in the community? What about socioeconomic conditions in the community?
How would you describe where your students see themselves headed after high school?
How would you describe the relationship between the schools and families in your district? On a scale of 1-10, with 1 being a total lack of trust and 10 being 100% trusting between parents and teachers, how would you rate the trust between schools and families? Why did you choose that number?

Implementation Status:

1. How would you describe PBE to different members of your community?

Probe:

How would you describe the ideal PBE to a student? a parent? A new teacher who was not familiar with the work you are doing?
If I were to come into the high school today and ask a teacher/student/parent to describe proficiency-based education, what would that person say?

2. When and where did the efforts to implement proficiency-based education begin (what year and with which teachers/school)?

Probe:

What did you do first? Why?
Where are you now?
What's next?

How did the staff respond initially when your implementation efforts first began? Has staff response changed over time?

3. Who are the biggest supporters or champions of the work? Why do you think that has been the case? Who has been hardest to convince? Why?

4. In your experience, has implementation of PBE been easier with some student groups?

Probe:

Our funder, Jobs for the Future, uses the term "Student-Centered Learning" to refer to a range of strategies, of which proficiency-based education is a component. How would you understand PBE as a student-centered approach?
Do you think there are some students who benefit from it more than others? Why?

Barriers and Facilitators:

5. I'd like to talk about some of the different elements that are related to the idea of PBE and get your input about what you see as both barriers to and facilitators for implementing each element at your school.

For each of the terms I am going to present, can you first provide me with a description of what that means/looks like in your school and then second, what has been challenging and what has been helpful in terms of implementing this element:

- Personalization.
- Opportunities to progress through demonstration of mastery or proficiency.
- Development of students' social and emotional competencies (voice, ownership, growth mindset, etc.).
- Opportunities to demonstrate learning beyond the school setting.
- PB diploma system.
- PB grading.

Element	Barriers	Facilitators
Personalization		
Opportunities to progress through demonstration of mastery or proficiency		
Development of students' social and emotional competencies		

Opportunities to demonstrate learning beyond the school setting		
PB diploma system		
PB grading		

6. What is it about your own school/district (context/environment) that supports this reform and what is it about the school/district that might make this reform somewhat more difficult?

7. How will you know whether you're successful in implementing PBE?

Parent Focus Group Protocol

Context (context: school)	1. Tell us about your children who are attending this school. Probes: *What classes do they most enjoy?* *What kinds of things are they involved in?* *Do they work?* *What is their attitude about school?*
Context (context re: understanding PBE)	2. You probably know that the school/school district is doing something different. That is: Student progress is being evaluated somewhat differently and students may be moving through their work at different paces, even in the same classes. Does this sound familiar? Are you familiar with the term proficiency-based education or PBE [note: modify terminology based on information from principal interview]? How do you know about this term and what do you know about it? Probe: *What do you think the reasons are that the state of Maine, and the school, is implementing PBE?*
Implementation (implementation status student level)	3. Do you know whether your child is experiencing PBE? If so, have you noticed any difference in the things your children are asked to do?
Implementation (supporters)	4. What do you see when your child(ren) comes home from school that is different from the past (homework, grades, etc.)? What do you think about it? Do you think these changes that we've discussed relate to PBE benefit or could benefit your child(ren)? Why or why not?
Implementation (implementation: groups of students)	5. What do you see as the advantages and disadvantages of the changes brought on with PBE (for all students)? What are you excited about for your child related to PBE? What worries you?

Implementation (implementation: measuring success)	6. What would it mean for your child(ren) to be successful in a PBE classroom or school? What would it look like for the school to be successful at making these changes?

Student Focus Group Protocol

Context (context: school)	1. Tell us about the kids in this school. Probes: What activities do they get involved in? Do they do after-school activities? Do they work outside of school or have family responsibilities?
Context (context re: understanding PBE)	2. You probably know that the school/school district is doing something different. That is, student progress is being evaluated somewhat differently and students may be moving through their work at different paces, even in the same classes. Does this sound familiar? Are you familiar with the term proficiency-based education or PBE [note: modify terminology based on information from principal interview]? How do you know about this term and what do you know about it? Probe: What do you think the reasons are that the state of Maine, and the school, is implementing PBE?
Implementation (implementation status: ckassroom/student level)	3. Do you have PBE in any of the classes you take? What does it look like? How does PBE in a class make it different from other current or past classes you've taken? Is anything different this year from the past? Probe: Do you see some teachers or subjects in which teachers are doing more with PBE than others? Why might that be?
Implementation (implementation: supporters)	4. Do you feel like you benefit from, or could benefit from, PBE, or the changes we've discussed, in your classes? Why or why not?
Implementation (implementation: groups of students)	5. Do you think there are some kids who benefit the most or the least from PBE? (Asked another way: Are there kids for whom these changes are really helpful and/or some for whom the changes are a problem? Why? Why not?)

Implementation (implementation: measuring success)	6. What would it mean for students, and you in particular, to be successful in a PBE classroom or school?

Teacher Interview Protocol

Context (context: classroom and school)	1. Tell me about the makeup of the students in your classes this year. Probes: *What are they like academically?* *Are they involved in school activities? What type?* *Do they have work or family responsibilities outside of school*? *Do the students in your classes differ from the students in the school overall*?
Context (context re: understanding PBE)	2. How do you describe proficiency-based education (PBE)[note: modify terminology based on information from principal interview] to your students? Parents? Probe: *What do you think the reasons are that the state of Maine, and the school, is implementing PBE?*
Implementation (implementation: status classroom level)	3. I am going to list four different aspects of student centered learning that are related to PBE and I want you to think about each one in turn. For each, can you tell me what, if anything, you are doing in your classrooms related to this component? How, if at all, has it changed what you're doing in the classroom? • Personalization • Opportunities to progress through demonstration of mastery or proficiency (relatedly: multiple assessment?) • Development of students' social and emotional competencies (voice, ownership, growth mindset, etc.)

	• PBE grading
Implementation (implementation: status classroom level)	4. Let's consider those four elements again. So far, what has been most successful and most challenging in implementing PBE at the classroom level? • Personalization • Opportunities to progress through demonstration of mastery or proficiency • Development of students' social and emotional competencies (voice, ownership, growth mindset, etc.) • PBE grading Probe: *How has implementation varied by different grade levels and tracks (e.g., CTE and online course takers)? Are different students having different experiences with the different aspects of SCL?*
Implementation (implementation: supporters)	5. Thinking about your experience with PBE to date, on a scale of 1 to 10, how would you rate your support of PBE, with 1 being "very skeptical and uncertain of its value" and 10 being "really enthusiastic and invested?" Why did you assign yourself that number? Probe: *Are there individuals or groups of teachers whom you would describe as having differing levels of enthusiasm for and investment in implementation of PBE? Do you see that changing over time?*
Implementation (implementation: groups of students)	6. In your experience, has implementation of PBE been easier with some groups or types of students than others? If so, which students, and why? Probe: *Do you see certain groups of students, such as those with IEPs or high-performing or CTE students, benefiting more or less from PBE?* *Do you think it's equally important for all students?*
Implementation (implementation: leadership)	7. Let's talk about leadership. What do you see as the principal's role in moving PBE forward? What do you see as your role at the classroom or school level in moving PBE forward?

Implementation (implementation: measuring success)	8. How will you know if you're successful in implementing PBE in your classroom? Is success in PBE the same for all students or all classrooms? If not, how does success look different?
ADDITIONAL QUESTION	9. In our analyses of the student survey data, we are uncovering some patterns in the way students are responding to the items on the survey. In fact, we see as much variation within schools as across them, with the responses clustering into three groups, which might roughly be described as high, medium, and low—the high response group are those that report having more experiences with being able to demonstrate their learning in a range of ways or being able to work in small groups, or get credit for project work, for example. These groups are similar across all the schools, despite differences in how each school may be implementing PBE. So, we're wondering why you think we might be seeing kids responding in these different ways. To what would you attribute these different groupings?
GIFT CARD	We would like to give you a $25 Amazon.com gift card to thank you for your participation. Please sign this receipt for our accounting department records.

Principal Interview Protocol

Context (contex: community)	1. Tell me about the makeup of your student body and the community that your high school serves.
Context (context re: understanding PBE)	2. Is proficiency-based education (PBE) the terminology you use to describe the reform work you are doing in the school? What do you think the reasons are behind the state's implementation of the PBE requirement? How would you describe PBE to different members of your community? Probe: *Teachers? Students? Parents? Other community members?*
Implementation (implementation status: school level)	3. I am going to list four different aspects of student-centered learning that are related to PBE, and I want you to think about each one in turn. For each, can you tell me what, if anything, the school is doing to implement this component? How, if at all, has it changed what's happening in schools? • Personalization • Opportunities to progress through demonstration of mastery or proficiency (relatedly: multiple assessment?) • Development of students' social and emotional competencies (voice, ownership, growth mindset, etc.) • PBE grading
Implementation (implementation: status school level)	4. Let's consider those four elements again. So far, what has been most successful and most challenging in implementing PBE? • Personalization • Opportunities to progress through demonstration of mastery or proficiency (relatedly: multiple assessment?) • Development of students' social and emotional competencies (voice, ownership, growth mindset, etc.) • PBE grading Probe: *How has PBE implementation unfolded in your school so far? Where are you and what's next?*

	How has implementation varied by different grade levels and tracks (e.g., CTE and online course takers)? Are different students having different experiences with the different aspects of SCL?
Implementation (implementation: supporters)	5. Who are the biggest supporters or champions of the work? Why do you think that has been the case? Who has been hardest to convince of its value? Why? Probe: *Are there groups of teachers (by grade level, department, tenure, for example) whom you would describe as having differing levels of enthusiasm for and investment in implementation of PBE? Do you see that changing over time?*
Implementation (implementation: groups of students)	6. In your experience, has implementation of PBE been easier with some groups or types of students than with others? If so, which students, and why? Probe: *Do you see certain groups of students benefiting more or less from PBE? Do you think it's equally important for all students?*
Implementation (implementation: leadership)	7. Let's talk about leadership. As principal, what do you see as your role in moving PBE forward? What do you see as the teachers' role?
Implementation (implementation: measuring success)	8. How will you know whether you're successful in implementing PBE at the school level?
Additional Question	9. In our analyses of the student survey data, we are uncovering some patterns in the way in which students are responding to the items on the survey. In fact, we see as much variation within schools as across them, with the responses clustering into three groups, and these groups might roughly be described as high, medium, and low—the high response group are those that report having more experiences with being able to demonstrate their learning a range of ways or being able to work in small groups, or get credit for project work, for example. These groups are similar across all the schools, despite differences in how each school may be implementing PBE. So, we're wondering why you think we might be seeing kids responding in these different ways. To what would you attribute these different groupings of kids?

References

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. Structural Equation Modeling: A Multidisciplinary Journal, 21(3), 329-341.

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.; pp.509-535). Thousand Oaks, CA: Sage.

Competency Works. What is Competency Education? Briefing paper, September 2014. https://www.competencyworks.org

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes.

Dziak, J. J., Bray, B. C., Zhang, J., Zhang, M., & Lanza, S. T. (2016). Comparing the performance of improved classify-analyze approaches for distal outcomes in latent profile analysis. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *12*(4), 107.

Finn, J. D. (1989, Summer). Withdrawing from School. *Review of Educational Research*, *59*(2), 117-142.

Fixsen, D.L., Naoom, S.F., Blasé, K.A., Friedman, R.M., Wallace, F. (2005). Implementation Research: A Synthesis of the Literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, *27*(15), 2865-2873.

Glazer, B. C., & Strauss, A. L. (1967). The discovery of ground theory.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Jobs for the Future. (2014). *Putting students at the center: A reference guide*. Boston, MA.

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and personality psychology compass*, *2*(1), 302-317.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, 175-198.

McLachlan, G., & Peel, D. (2000). Mixtures of factor analyzers. *Finite Mixture Models*, 238-256.

Miles, M. B., Huberman, A. M., Huberman, M. A., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. sage.

Miller, L. (2001). School-university partnership as a venue for professional development. *Teachers caught in the action: Professional development that matters*, 102-117.

Muthén, L. K., & Muthén, B. (2015). Mplus. *The comprehensive modelling program for applied researchers: user's guide*, *5*.

Nagin, D. (2005). *Group-based modeling of development*. Harvard University Press.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling*, *14*(4), 535-569.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, 111-163.

Ryan, S., & Cox, J. D. (2016). Guide to the Competency-Based Learning Survey for Students. REL 2016-165. *Regional Educational Laboratory Northeast & Islands*.

Steele, J. L., Lewis, M. W., Santibanez, L., Faxon-Mills, S., Rudnick, M., Stecher, B. M., & Hamilton, L. S. (2014). Competency-Based Education in Three Pilot Programs.

Stump PhD, E., Silvernail PhD, D., Fallona PhD, C., & Moran Gunn, L. (2013). Maine's improving schools: An examination of distinguishing features of a sample of Maine's improving public schools.

Stump, E. K., & Silvernail, D.L. (2014). *Implementation of a proficiency-based diploma system: Early experiences in Maine*. Gorham, ME: University of Southern Maine.

Stump, E. K., Doykos PhD, B., & Fink, P. (2017). Proficiency-based High School Diploma Systems in Maine: Implications for College and Career Access.

Sturgis, C., & Patrick, S. (2010). *When success is the only option: Designing competency-based pathways for next generation learning*. Vienna, VA: International Association for K-12 Online Learning. Retrieved from http://www.inacol.org/research/docs/iNACOL_SuccessOnlyOptn.pdf

Sturgis, C., Chugach School District: A Personalized, Performance-BasedSystem, International Association for K-12 Online Learning, 2016. Content

Team, R. C. (2015). R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2014.

Torres, A. S., Brett, J., & Cox, J. (2015). Competency-Based Learning: Definitions, Policies, and Implementation. *Regional Educational Laboratory Northeast & Islands*.

Voelkl, K. E. (1996). Measuring students' identification with school. *Educational and psychological measurement*, *56*(5), 760-770.

Voelkl, K. E. (1997). Identification with school. *American Journal of Education*, *105*(3), 294-318.

Wolfe, R. E., Steinberg, A., & Hoffman, N. (2013). *Anytime, Anywhere: Student-Centered Learning for Schools and Teachers*. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138.