CADRE

SEPTEMBER 2024

Learnings from a Large District's Experience with Using Performance-based Assessments to Evaluate Graduation Competencies

Elena Diaz-Bilello, Kaitlin Nath, Adam York

A report prepared by the Center for Assessment, Design, Research and Evaluation (CADRE) at the CU Boulder School of Education.

School of Education
UNIVERSITY OF COLORADO **BOULDER**

Acknowledgements

We are deeply grateful to our state and district partners who made this study possible. We extend special thanks to Angela Landrum at the Colorado Department of Education for the continued partnership and her valuable feedback on this report.

About CADRE

The Center for Assessment, Design, Research and Evaluation (CADRE) is housed in the School of Education at the University of Colorado Boulder. The mission of CADRE is to produce generalizable knowledge that improves the ability to assess student learning and to evaluate programs and methods that may have an effect on this learning. Projects undertaken by CADRE staff represent a collaboration with the ongoing activities in the School of Education, the University, and the broader national and international community of scholars and stakeholders involved in educational assessment and evaluation.

Suggested Citation

Diaz-Bilello, E., Nath, K., & York, A. (2024). Learnings from a Large District's Experience with Using Performance-based Assessments to Evaluate Graduation Competencies. Boulder, CO: The Center for Assessment, Design, Research and Evaluation (CADRE), University of Colorado Boulder.

Please direct any questions about this project to:

elena.diazb@colorado.edu

Executive Summary

This study conducted by the Center for Assessment, Design, Research and Evaluation (CADRE) at the University of Colorado Boulder documents the work undertaken by a large metro school district in Colorado to use collaboratively-developed performance-based assessments (PBAs) to evaluate graduation readiness or competencies. Key learnings from this work include:

- Despite the time-intensive work involved with enacting PBAs, teachers implementing this work see value in using PBAs since student thinking and reasoning are made more visible and teachers can use this information to better inform instructional steps.
- Expectations for grading and evaluating graduation competencies using the PBAs vary across teachers. This highlights the importance of ensuring that shared expectations or standards for this work are discussed and normed across teachers and schools.
- Teacher leaders at the schools shared that in sharp contrast to multiple-choice tests, the PBAs provide more expansive and fair opportunities for minoritized groups to demonstrate what they know and can do through the various modalities offered to all students to engage with the tasks.

Due to the large size of this district and the site-based management approach taken by district schools, the district is currently navigating tensions to ensure that shared expectations on graduation competencies using the PBAs can be achieved across schools. For this district, norming conversations will take place through professional learning opportunities held during the upcoming school year to ensure expectations are communicated clearly to both students and teachers across district schools.

Table of Contents

Introduction

The 2021-22 school year marked the full implementation year for Colorado's graduation guidelines. The Colorado Department of Education (CDE) developed the guidelines in consultation with stakeholders to ensure that high school students had different options to demonstrate readiness for entry into postsecondary or workforce opportunities. One of the graduation menu options, collaboratively-developed, standards-based performance assessments, was introduced along with the capstone/portfolio approach, to provide more authentic and engaging demonstrations of learning for high school students.

This report developed by the Center for Assessment, Design, Research and Evaluation (CADRE) at the University of Colorado Boulder documents learnings from work taking place in one large metro school district to use collaboratively-developed performance-based assessments (PBA) for evaluating graduation competencies in English Language Arts (ELA) and math. The findings in this report come from two high schools that have a longer history of using the PBAs relative to other schools in the district.

Background

At the height of the pandemic disruptions in 2020-21, the Ginkgo School District[1] (GSD) introduced common performance assessments as an option for students to engage in credit recovery and to demonstrate readiness for graduation. Teachers and instructional coaches at GSD high schools developed these performance-based assessments (PBAs) to align with the state's content standards and essential skills. For example, a math task developed by teachers invited students to apply mathematical knowledge and skills to real-world scenarios such as applying financial literacy skills to manage household finances. These tasks were not designed for a specific curriculum, and the district provided a recommended scope and sequence that could be followed prior to administering a given PBA.

According to district administrators, the PBAs originally served two purposes. The primary purpose was to broaden pathways for students to engage in meaningful demonstrations of learning outside of standardized testing approaches. The secondary purpose was to expand opportunities for students to engage in credit recovery due to course work missed with school closures and other disruptions encountered in the home lives of students during the height of the COVID-19 pandemic.

In the 2021-22 school year, GSD administrators communicated the policy to all high schools that the PBAs could no longer be used for credit recovery purposes, and that they could only be used to determine whether students met graduation competency requirements in ELA and math. This policy was enacted since district administrators and the PBA developers who supported this new policy direction noted that having the PBAs serve two distinct purposes generated confusion for educators and school leaders. They also observed that many schools were defaulting to using the PBAs for credit recovery rather than for graduation competency.

[1]The Ginkgo School District and other names that appear in this report are pseudonyms assigned to protect the confidentiality of the district and school participants.

Currently, all traditional high schools in the district administer the PBAs to all students. Although only a select number of PBAs[2] can be used to evaluate graduation competency, most students fulfil competency through test-based options available on the menu (e.g., meeting benchmark scores on the SATs or the ACT WorkKeys). Each school determines the policy for grading and evaluating competency on the PBAs. Knowing that grading and competency policies differ across sites, we engaged two GSD high schools in a case study to understand the processes and decisions used by their teachers to establish these two key decision points for students in math. This work was focused on math since this disciplinary area is well-known to be challenging for teachers to enact and/or develop authentic PBAs (Turner et al., 2009; McREL, 2010; Showalter, 2013). We also conducted descriptive analyses of the PBA outcomes at the same case study schools to gain an initial sense of how these results compared to other outcomes listed on the graduation menu options. These exploratory analyses were completed to learn the extent to which PBA results provide discrepant or similar signals to other graduation menu options in the 2022-23 school year.

The following three questions guided the case study work at the two sites:

Guiding Question 1 (GQ1): In what ways are the PBAs providing teachers with valuable information about what students know and can do on targeted math knowledge and skills?

Guiding Question 2 (GQ2): How are schools using the results from the math PBAs to help determine graduation competency for students in this disciplinary area?

Guiding Question 3 (GQ3): How do the results from the PBAs compare with results from other benchmarks used in other options on the graduation menu?

In the following section we address the sample and methods used to address each of these three questions.

Defining the Sample and Methods

The data used to address all three questions were restricted to two high schools: Pinyon High School (PHS) and Engelmann High School (EHS). Similar to the student population served at other schools throughout this metro district, over 70 percent of students enrolled in both schools belong to minoritized student groups. These schools were selected for two reasons: according to district staff, these schools have more experience integrating PBAs in professional learning community (PLCs) discussions compared to other traditional schools in GSD, and these two schools also have math teacher leaders or instructional coaches who helped develop the math PBAs.

To address GQ1 and GQ2, we observed five PLCs (two at PHS and three at EHS) focused on discussions involving the PBAs, and interviewed teacher leaders or the PBA leads steering the PLC work. We carried out interviews with the PBA leads to clarify questions we had following observations of PLC activities, and to understand the policy used by the PLC teachers to

[2]Two PBAs were developed for each high school grade (9-12). The district has identified specific PBAs beginning with grade 10 that can only be used to evaluate students for competency, and schools can further identify which of these competency-PBAs will be used to evaluate graduation competency.

establish graduation competency for math. We also interviewed the district math coordinator who accompanied us on the site visits to understand district policy and positions with the PBAs, and triangulated findings from the observations with her.

For GQ3, the GSD assessment office provided student-level demographic and performance data for the 2020-21 through the 2022-23 academic years. The PBA data were restricted to two disciplinary areas evaluated in the expanded graduation menu (ELA and math). These data were further restricted to grades 11 and 12 to reflect the PBAs commonly used by the schools to evaluate graduation competencies. The data indicate whether students received an A, B, or C on the PBAs, or whether they took the PBAs but received lower than a passing grade. In addition to PBA data, the district provided student-level data on several other outcome variables including PSAT/SAT performance, average daily attendance (ADA), AP test scores, IB test scores, GPA, behavior data, performance in concurrent enrollment courses, and post-high school plans. Table 1 displays the number of students and the demographic composition of students in the initial sample (in grades 11 and 12) by school and year.

Table 1. Number of students in grades 11 and 12 and demographic composition by school and year

School	Year	N	% FRL	% ELL	% IEP	% Minority	% Female
EHS	2021	802	0.69	0.74	0.13	0.91	0.52
EHS	2022	765	0.75	0.75	0.13	0.91	0.50
EHS	2023	678	0.80	0.79	0.13	0.92	0.51
PHS	2021	946	0.45	0.42	0.09	0.72	0.48
PHS	2022	855	0.51	0.40	0.09	0.72	0.49
PHS	2023	818	0.57	0.42	0.10	0.77	0.49

The table in Appendix A displays descriptive statistics for each of the variables of interest, including the number of students for whom data are available and included in the sample. In total, the initial sample includes 9990 observations across 3504 distinct students. While all students in the sample have ADA data and GPA data, far fewer students have AP, IB, concurrent enrollment, and PBA data. In our results section, we provide supportive detail to justify our decision to restrict the analytical sample to the 2022-23 school year.

Analytical Approach

Qualitative Data Analyses

To analyze the qualitative data, we employed a general inductive approach (Thomas, 2006) with an eye toward addressing the first and second guiding questions. An inductive approach allows findings to emerge from multiple, close readings of raw data with the goal of identifying frequently occurring or important themes that have connections to or can be supported by extant literature.

We first reviewed and selected excerpts from our observation field notes that contributed insights for our guiding questions. We then reviewed and initially coded all interview transcripts following two separate readings. The initial codes consisted of broad categories that included: goals, challenges, solutions, variability, and consistency. Following the first application of general categories, we established a second layer of codes to describe frequently occurring phenomena or issues related to these general categories. For example, under the broad category of "challenges", the review of data pointed to a frequently occurring observation about the "lack of common criteria" to guide grading or competency decisions on the PBAs as captured by the different sources of data (i.e., field notes and interviews). This sub-category along with other co-occurrences coded such as "year to year criteria changes" generated an important theme around the lack of consistent criteria or common policies driving competency and grading decisions. Following a review of the coded excerpts, we culled representative interview and observation data for inclusion in the presentation of findings.

Descriptive Analyses

To address GQ3, we first examined the distribution of performance on the PBAs for the two schools. We restricted the analyses to students who passed the PBAs and examined the extent to which this subset of students passed the benchmarks for other available graduation menu outcomes. We also examined how each passing grade on the PBAs is associated with cumulative GPAs and with chronic absenteeism rates.

Results

Moving into our results section, we begin with findings from the site visits and interviews before moving into the preliminary results from examining PBA outcomes.

Using PBAs for Evaluating Graduation Competency

At both high school sites, the PLCs were focused on discussing the same grade 11 roller coaster PBA. For this multi-part PBA, students designed a roller coaster using a set of equations that included factoring in other variables such as length of time and distance. The task required students to work in collaborative groups as well as to reflect independently on their understanding of the task demands. Students were also asked to consider essential skills such as how well they contributed to their team's success following the completion of the PBA.

In our initial site visit to both schools, teachers had not administered the PBA and were in the process of adjusting the task and the rubric used to evaluate student work. For the second PLC visit at both schools, the agenda indicated that teachers would engage in student work analysis to evaluate PBA outcomes after administering the four-day task. A third site visit took place at EHS since the math team planned an additional PLC meeting to analyze student work relative to graduation competency decisions.

Although our site visits were limited in scope (i.e., two visits to PHS and three visits to EHS), the case study activities revealed promising aspects of how teachers were implementing the PBAs to learn about what students know and can do on the targeted math knowledge and skills; as well as surfaced challenges with the ongoing PBA implementation work. We first highlight findings associated with promising aspects of the PBA implementation work, before elaborating on areas that we identified as challenges. In a few instances, we address findings restricted to just one school site, if other sources (i.e., observations, PLC lead interview, and/or district interview) point to how a specific finding appears to be systemic within the school and potentially at other schools outside of our case study work.

Promising PBA Implementation Work

Successful Curriculum Integration

During our observations, the discussions in the PLCs observed did not surface any challenges experienced by teachers to fit the PBA into their scope and sequence. In both schools, the PLC teachers coordinated the timing and administration of this PBA and used the PBAs as replacement tasks for the typical assessments that they would use in their curriculum. In other words, teachers in these PLCs did not appear to struggle with an issue raised last year by the PBA task developers, about difficulties they observed with colleagues who were trying to embed the task in the math curriculum.

According to the PLC lead in PHS, she noted that the PBAs selected for evaluating graduation competency were a "good fit" for the district curriculum and the teachers in her school. She also mentioned that the teachers in her school did not see these PBAs as being "more different" than

the assessments that they would typically use at their school to evaluate student learning. For PHS in general, the teachers have found these assessments to cohere with their instructional model. This finding shared by the PLC lead at PHS is promising since at a national level, teachers have reported perceiving limited coherence in their instructional system comprised of key components supporting instruction such as curricula, professional development and assessments (Pauketat et al., 2023).

In contrast to the sentiments shared by the teacher leader at PHS, the PLC lead at EHS noted that her specific math team was the only team focused on the PBAs because the ELA and other math teams were more focused on defining success solely through tests. The PLC lead at EHS shared that, "even though my team knows that [students] in our classes don't have to do so well on SAT…the reality is everybody else [in the school] is focused on that." According to the EHS PLC lead, her math team is "very separated from the math department," since the priority of her team is not to focus on standardized test results, but to have students engage in tasks that, "allow for student discourse, student thinking and [different opportunities for] communication of thinking." A similar reflection shared in this interview surfaced in a PLC discussion at EHS where teachers shared concerns about their school leaders wanting to see more traditional and didactic instruction taking place in classrooms rather than expressing support for the type of collaborative learning activities emphasized in the PBA tasks. For this math team at EHS, the instructional incoherence they experienced was coming from their school leaders' expectations to have teachers conform to an instructional model and assessments that were incompatible with their desire to focus on more project-based and authentic assessment work.

Expanding Equitable Assessment Opportunities through PBAs

Teacher discussions about the PBAs at both sites and the interview data from the PLC leads suggests that the PBAs provided more equitable opportunities for students to demonstrate their learning compared to more traditional models of assessments. For example, during the PLC discussions at EHS, the teachers noted that the different modalities for showing student work (i.e., written reflections, modeling designs and equations on Desmos and paper, creating posters, etc.) provided deeper ways for Multilingual Learners (MLs) to demonstrate their understanding of content compared to traditional multiple-choice tests.

In our interview with the PLC lead at PHS, she summarized that PBAs are accessible to most students, including those with special needs, and gave them broader opportunities to demonstrate their understanding. When reflecting on student work from the PBAs, she remarked that students, "were able to [collaborate] together, and even with language barriers, show a greater understanding and contribute to the learning [through the collaborations]." The PLC lead at EHS communicated a similar sentiment and shared that the PBA, "is a better way to see what [students] actually know and how they can communicate [and show their knowledge and skills]." She shared that the value of the PBAs was especially apparent in cases where students struggled on traditional tests, and benefitted from an opportunity to, "communicate or show their work differently." She also discussed how the PBAs are a valuable instructional tool because the project format is memorable for students and something that is easy to return to as a reference point in future lessons.

The notion that the PBAs made student thinking and understanding more visible to teachers surfaced in our observation of the one PLC meeting at EHS focused on examining student

work. In this PLC session, the teacher leader brought student work products from two of her classes. She structured an activity for the other two grade 11 math teachers to review students' final posters and to use the rubric to rate student work. The teachers discussed the individual posters and the rationale for their ratings based on the categories listed in the rubric, and compared ratings across posters.

After the initial ratings discussion, the teacher leader read aloud selected student written reflections with the group. This was followed by a discussion of how the group might change their ratings when factoring in student thinking from these self-assessments. For example, in one case the teachers discussed changing a rating from a 2 (not achieving graduation competency) to a 3 (achieving graduation competency) based on the written reflections the student provided which surfaced deeper conceptual understandings of the algebraic functions explored in the task. According to the PLC leader, the written reflections provided another avenue for her to evaluate student work. She noted that the reflections also aligned with the competency descriptor under Level 3 in the rubric that required students to, "effectively communicate their thinking and their understanding and make sense of [the task demands]." This line of reasoning then compelled her colleague who initially rated the student work at a Level 2, to factor in responses from the self-assessment and move the rating up to a Level 3.

Consistent with literature focused on elevating important culturally responsive classroom strategies (Taylor, 2022; Evans, 2023), the roller coaster PBA discussed in the PLCs at EHS and PHS was designed to appeal to diverse learners by encouraging students to engage in group collaborative work, and by providing multiple modalities to demonstrate mathematical knowledge and skills. As indicated by the student work analysis discussion and the interviews with the PLC leads, these opportunities allow teachers to evaluate graduation competency based on a broader set of evidence compared to limited test-based approaches. This finding is especially salient since providing students with various opportunities to demonstrate knowledge and skills through authentically designed PBAs may help mitigate possible teacher biases for viewing and assessing student work (Chism, 2022; Darling-Hammond, 1994; Solano-Flores & Li, 2009). Though not observed at the PLC discussions at PHS and EHS, the district math coordinator shared that some secondary math teachers in other PLCs observed across the district have explicitly stated that they do not believe their students are capable of the level of rigor required by the mathematics curriculum. And as observed in the third PLC at EHS, unpacking student knowledge and skills through a variety of task demands from the PBA provided opportunities for those teachers to pay more attention to student strengths and assets by broadening the evidence-base used to evaluate what students know and can do. In other words, the different modalities allowed all of their students to demonstrate that they were capable of meeting the rigor of the math curriculum.

Challenging PBA Implementation Work

Lack of Shared Criteria for Evaluating Graduation Competency

A surprising finding that surfaced from attending the PLCs at both schools is that the PLC teachers had not finalized criteria for determining competency decisions, and that discussions were still ongoing on this matter. One exchange between teachers during the first PLC visit at PHS highlighted the unresolved nature of establishing competency in math with the PBAs. One teacher asked the question, "If a student is able to find the correct answers with help of others,

is it competency?" A second teacher replied, "If they are really working together in a group, they should get the grade." The original teacher continued her line of questioning, "But does that mean proficiency?" A third teacher, responded, "Some shouldn't pass based on their lack of individual effort." The PLC lead teacher concluded, "It is okay to differentiate. This means some in the group don't get as high of a grade. [It is] better to base competency on [the] individual body of evidence." During the second PLC discussion at PHS, the district math coordinator asked the PLC teachers if the 12th grade PBA could be used to evaluate competency if a student did not pass the grade 11 PBAs. A teacher responded that this would depend based on the "counselor's decision."

Similarly, the discussions at EHS also revealed that no clear policy on evaluating graduation competency using the PBAs had been established at the school. At EHS, the three math teachers indicated in the first PLC meeting that conversations about graduation competency would take place later in the school year (early May). The PLC lead and one other teacher shared that in the past, the school used different rubrics to evaluate competency but would return to the "CDE rubric" or the performance level descriptors used for the state's summative assessment, to evaluate student work for competency purposes. During the third PLC site visit, the activities observed focused on using the CDE rubric more consistently to determine if a student's work product from the PBA met graduation competency expectations for math. However, despite the norming activity observed at the school, the PLC lead indicated in a separate interview that ultimately, "there are no [school-based] policies for competency…It's what I decide."

At the second PLC meeting observed at PHS, teachers planned to conduct student work analysis. Instead of focusing on student work, the teachers reflected on the overall successes and challenges experienced in administering the PBA. When we asked the PHS teacher leader about their typical practices with student work analysis, she shared that PLC work focused on student work analysis to calibrate grading "varies," and depends largely on "where it falls in the teacher calendar and what [we] have time for." Additionally, she mentioned that the PLC teachers did not have a common protocol to use for evaluating student work. This response from the PLC lead coupled with the lack of shared protocols to guide student work analysis activities, strongly suggests that clear success criteria for grading which leads into graduation competency had yet to be solidified for this group of PLC teachers.

We consider the lack of a consistent standard for evaluating competency to be unexpected since teachers at these schools have administered the PBAs for three consecutive years. However, the PLC discussions at both sites and interview data from the two PLC leads suggest that there are no shared or common criteria established in the school for how students can achieve competency on these PBAs for graduation purposes. The interview with the district math coordinator also revealed that the lack of shared criteria expands beyond these two schools since this appears to be an area that all schools are trying to "figure out." As shared by the district math coordinator, "I don't think we have had any type of district wide conversation and therefore the definition [for graduation competency using PBAs] doesn't exist or it doesn't at least align in multiple spaces."

Although the district intentionally devolved this competency decision to school sites to uphold a policy of site-based decision making, the fact that the success criteria for establishing competency using the PBAs are not currently clear to all teachers in each site, would by

extension suggest that the criteria are not clear to students. We acknowledge the importance of providing teachers with agency to establish competency criteria, however, any definition developed should ideally be co-designed across teachers to ensure common criteria are established across classrooms; and, subsequently clearly communicated to students. In other words, the current lack of transparency on defining competency using the PBAs within each site is problematic since this implies the lack of a fair and transparent standard established for meeting graduation competency on these assessments. This lack of a shared standard across teachers for this important use case (i.e., graduation competency) can potentially contribute to inequitable grading practices and policies for students (Feldman, 2019a, 2019b).

Continuous PBA Modifications and Scaffolding

In the PLCs we observed, teachers expressed that the roller coaster PBA represented one of the more rigorous math PBA that met what they believed were their individual expectations for demonstrating graduation competency in the targeted math knowledge and skills. The PLC lead from PHS shared that in the initial roll out of the PBAs, "the earlier version of the [roller coaster PBA] did not have enough teacher input to be …rigorous enough to hit [the expectations] we wanted." According to the PLC lead, teacher input was not solicited in the first two years due to the COVID-19 disruptions and the immediate need to use the PBAs to evaluate student learning during this period. The task was revised over time by the PLC to, "make sure that we were really hitting the state standards at the 11th grade level."

Yet despite having identified this as an appropriate task to include in the evaluation of graduation competency for math, the PLC activities observed continued to make modifications to the task with a focus on clarifying task demands for students and teachers. For example, at PHS, the teachers developed a guiding worksheet to indicate which parts of the assignment should be completed by an individual student versus a group, and established more clarity to the sequence of activities pursued each day. According to one teacher, the goal of these changes, "[is to try] to be clearer about what [students need to accomplish] on day one versus day two, etc." Other teachers agreed in the session that these updates would be helpful to make the task demands clearer to students, which would subsequently help students prioritize and complete activities on each day. The PLC also focused on modifying the rubric to ensure that success criteria were clearly communicated to establish expectations for students, and to inform grading for teachers. Similar to the PHS group, the teachers at EHS also used a portion of the PLC time revising the rubric to establish clearer success criteria. They also adjusted one of the prompts to clarify the directions for students. Although the underlying construct evaluated remained the same, the modifications made at both schools suggested that key task features that directly influence outcomes (e.g., student directions, the rubric) still required refinements.

Additionally, in one PLC session observed at EHS, the issue of scaffolding surfaced, and the PLC lead shared that that she developed a supplementary reference sheet for MLs which included a list of math definitions and corresponding examples to use as models to guide their work. According to the PLC lead, she believed this reference sheet was instrumental for these students to complete the task. A scaffolding discussion also emerged in the PLC discussion at PHS. During the PLC observed, the PLC lead reflected with the group that in, "part one [of the task] the students did well, but on part two they were struggling with how to write the equation even though they have practiced this several times…" Another teacher concurred and noted, "[I] really have had to talk them through how to write the equation [and remind them] what the

roots are." In response, the PLC lead indicated taking a different approach, "[I] didn't want to walk them through [the equations], [I] want to be able to point to things they've done [in prior assignments], not walk them through it because it is a PBA."

In the case of the ML resource supplement, using this supplement while taking the PBAs could potentially lead to an over-scaffolded task - especially if these students are following the models provided in the supplement too closely rather than deriving an independent representation that reflects their own thinking. Additionally in the case of the two teachers discussing equations at PHS, this interaction surfaced that no clear consensus has been established in the PLC to determine how much scaffolding may be appropriate for students on the PBAs. Ideally, PLC discussions would address a range of possible strategies, such as engaging in more distributed approaches to ensure that the work is not over-supported, but discussions fell short due to limited time in each setting to evaluate the merits of scaffolds used. According to the district math coordinator, finding an ideal balance between maintaining the quality and rigor of the PBA and minimizing scaffolds is an area that other schools are also trying to find. The math coordinator noted that in other PLCs observed, she thought that that teachers might be scaffolding the PBAs to the point that these are becoming too easy for student and noted, "with a lot of the PLC work around modifications… potentially we might be over-scaffolding in certain ways to essentially check a box for students in the easiest way possible." If done well, scaffolds can be highly beneficial to students (Banse et al., 2016; Frederick et al., 2014). However, these high school PBAs are intended to increase student ownership over their learning, and this would require that teachers know when to engage in appropriate fading of scaffolds to move more control and responsibility to their students (Dove & Hollenbrands, 2014).

Findings from Preliminary PBA Data Collected

We begin this section with a broad overview of student performance on PBAs to explain our decision to first limit our descriptives to the 2022-23 school year, and then defer our analytical work for the upcoming 2023-24 school year.

Although we had data beginning in the 2020-21 school year, the PBA data from 2020-21 reflected outcomes that informed either competency or credit recovery. Since data collection at that time did not distinguish whether PBA grades recorded were used for one purpose or another, we could not distinguish which grades from 2020-21 were used for competency determinations. We then restricted the data to the 2021-22 and 2022-23 since 2021-22 marked the first year in the district that the PBAs could only be used for graduation competency purposes, and assumed that we could run descriptives using two years of PBA data. However, after encountering large inconsistencies in the data between the two years and by subject area, we further restricted our analyses to the 2022-23 year.

In PHS, for example, the competency data were inconsistently entered between the two subjects and years. According to the district, scores reported in math were documented as "pass/fail" in 2021-22 and entered as grades in 2022-23. In Table 2, the 334 students included in the PHS PBA Math (21-22) row are all students who passed the PBA; students who did not pass the PBAs are not included in the dataset. The different scoring rules used for the Math PBAs in PHS in 2022-23 means that the reported scores are not comparable across the two years. Additionally, according to the district, there was district-wide confusion regarding what constituted a passing grade on ELA PBAs in 2021-22. While the district intended for only grades

A through C to be considered passing, and thus only included grades A, B, and C into the learning management system, the district ultimately decided to include D as a passing grade. As a result, "NA" values on the PBA ELA tests in 2021-22 represented students who took the PBA and received a D.

Table 2. Number of observations by school, year, PBA subject

School	PBA test	A	B	C	NA	Total
EHS	PBA ELA (21-22)	25	13	8	10	56
EHS	PBA ELA (22-23)	2	4	6	NA	12
EHS	PBA Math (21-22)	18	26	10	NA	54
EHS	PBA Math (22-23)	44	52	57	NA	153
PHS	PBA ELA (21-22)	69	84	54	24	231
PHS	PBA ELA (22-23)	265	187	103	NA	555
PHS	PBA Math (21-22)	NA	NA	NA	334	334
PHS	PBA Math (22-23)	174	146	131	NA	451

Finally, we also learned that no data were entered for students that did not pass the PBA in 2023-24, and therefore we do not have an accurate tally of the total sample of students at both schools that completed the PBAs. Hence, the data in 2023-24 are incomplete and cannot be used to establish baseline information about PBA outcomes. Moving forward, we will coordinate with the district to ensure that we have data for all students that took the PBAs that count for graduation competencies to begin our baseline analyses of PBA outcomes in this school district.

Implications

A highly promising aspect of the PBA implementation work at these two schools is that the PLC teachers we observed believe that the performance-based demonstrations of learning using the district PBA tasks can provide more equitable learning opportunities for students with diverse backgrounds and learning needs compared to traditional multiple-choice tests. As a result, teachers gained better insights about their students from these higher-order tasks since these prompted them to make their thinking and reasoning visible. Additionally, despite the length of time (i.e., multiple class periods) required for these teachers to administer a PBA used for competency decisions, none of the teachers in the PLCs expressed complaints about the amount of time used to engage students with the PBAs. At the second PLC meeting observed in EHS, the teachers expressed a desire to further lengthen the PBA experience by potentially transforming the PBA into a larger capstone experience that would involve the external community as reviewers. The willingness for these teachers to use and embed the district PBAs in their curriculum is a strong indication that they see high value in investing their time and resources in engaging with the PBA work. This willingness is important since the burden of enactment and scoring demands associated with PBAs have historically deterred many districts and schools from engaging in this type of assessment work (Stecher, 2010).

Considering the larger size of this school district and the district's emphasis on allowing schools to engage in site-based decision making, it should not come as a surprise that grading and competency approaches taken using the PBAs varied within and across these two schools. However, if the district wants to ensure that there is a shared standard for what constitutes competency in Math and ELA across schools, then the results from this report point to the importance of calibrating expectations across schools. In other words, schools can establish their own processes, but should ideally share similar expectations and standards for student work that meets a passing grade and competency for math and other subjects. In our preliminary examination of PBA data collected in 2021-22 and 2022-23, the results from the two high schools show that the grading criteria changed each year as both schools tried different grading approaches and employed different benchmarks to communicate competency outcomes. The observations of PLCs held at the two schools also revealed that teachers have not yet formalized processes on grading the PBAs and evaluating competency.

Although establishing a shared competency standard appears to be a work in progress for both sites, this presents an opportunity for the district to engage with all high schools to ensure that all site-based decisions adhere to a shared standard for competency. This would at the minimum, ensure that all sites have transparent and clearer consensus for competency, and can subsequently communicate this more clearly for all students. Providing opportunities for intra-school collaborations to norm expectations on competency may be helpful during districtwide professional learning days to ensure clearer criteria can be set across all schools. Another opportunity for district staff to consider for improving the ongoing district-wide PBA implementation work is to present different strategies and tools that can be used by teachers to scaffold the PBAs. Scaffolding can be structured and enacted through a variety of strategies, and teachers may find it helpful to draw on these resources so that they can avoid over-supporting and diminishing the rigor of enacted lessons, tasks, and the PBAs (Frey et al., 2023).

This type of support is particularly important since the level of scaffolding given has a large influence on outcomes achieved by students on the PBAs, and over-scaffolded tasks will not provide an accurate gauge of what students know and can do in these disciplinary areas.

Despite some of the ongoing challenges described in this report with the ongoing PBA work, we stress that it takes time to build the infrastructure and capacity to support complex reforms such as performance-based assessment work. That is, there is not a single pathway or template that works perfectly or can be easily adopted by a given site, due to different disciplinary and curricular demands, the size of a district, and a broad range of relational and contextual issues that vary across districts and schools (Cuban, 2021). Approaches to complex systems-wide reforms work better when districts take the time to adapt and demonstrate flexibility to meet their local and community context (Burns et al, 2020). In this district, the implications from the district's commitment to learn from the ongoing PBA implementation work at two high school sites are that that the findings documented in this report will be used to help inform future decisions and supports to strengthen the work over time.

References

Banse, H. W., Palacios, N. A., Merritt, E. G., & Rimm-Kaufman, S. E. (2016). Scaffolding English language learners' mathematical talk in the context of Calendar Math. *The Journal of Educational Research, 110*(2), 199–208. https://doi.org/10.1080/00220671.2015.1075187

Brown, B. (2019). *Science in the city: Culturally relevant STEM education.* Routledge.

Burns, D., Darling-Hammond, L., & Scott, C. (2019). Closing the opportunity gap: How positive outlier districts in California are pursuing equitable access to deeper learning. Palo Alto, CA: Learning Policy Institute.

Chism, D. (2022). *Leading your school toward equity: A practical framework for walking the talk. ASCD*, Arlington VA.

Darling-Hammond, L. (1994). Performance-based assessment and education- al equity. *Harvard Educational Review, 64*(1), 5–31. https://www.hepg.org/her-home/issues/harvard-educational-review-volume-64,-issue-1/herarticle/_351.

Dove, A., & Hollenbrands, K. (2013). Teachers' scaffolding of students' learning of geometry while using a dynamic geometry program. *International Journal of Mathematical Education in Science and Technology, 45*(5), 668–681. https://doi.org/10.1080/0020739X.2013.868540

Evans, C. M. (2023). Applying a Culturally Responsive Pedagogical Framework to Design and Evaluate Classroom Performance-Based Assessments in Hawai'i. *Applied Measurement in Education, 36*(3), 269–285.

Feldman, J. (2019a). *Grading for equity.* Corwin.

Feldman, J. (2019b). Beyond standards-based grading: Why equity must be part of grading reform. *Phi Delta Kappan, 100*(8), 52-55. https://kappanonline.org/standards-based-grading-equity-reform-feldman/

Frederick, M. L., Courtney, S., & Caniglia, J. (2014). With a Little Help from My Friends: Scaffolding Techniques in Problem Solving. *Investigations in Mathematics Learning, 7*(2), 21–32. https://doi.org/10.1080/24727466.2014.11790340

Gay, G. (2010). Acting on beliefs in teacher education for cultural diversity. *Journal of Teacher Education, 61*(1–2), 143–152. https://doi.org/10.1177/0022487109347320

Marion, S. (2018). The opportunities and challenges of a systems approach to assessment. *Educational Measurement: Issues and Practice, 37*(1), 45–48.

Pauketat, R., Wang E.L., Kaufman, J.H, Gittens, A.D. & Woo, A. Teachers' perceptions of coherence in K–12 English Language Arts and Mathematics instructional systems. Santa Monica, CA: RAND Corporation, 2023. https://www.rand.org/pubs/research_reports/RRA279-3.html.

Shepard, L. A., Penuel, W. R., & Pellegrino, J. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice, 37*(1), 21–34.

Solano-Flores, G., & Li, M. (2009). Generalizability of cognitive interview-based measures across cultural groups. *Educational Measurement: Issues and Practice, 28*(2), 9-18. https://doi.org/10.1111/j.1745-3992.2009.00143.x

Stecher, B. (2010). Performance Assessment in an Era of Standards-Based Educational Accountability. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Taylor, C. S. (2022). *Culturally and socially responsible assessment: Theory, research, and practice.* Teachers College Press.

Tzou, C., Meixi, Suárez, E., Bell, P., LaBonte, D., Starks, E., & Bang, M. (2019). Storywork in STEM-art: Making, materiality and robotics within everyday acts of indigenous presence and resurgence. *Cognition and Instruction, 37*(3), 306–326.

Appendix A

Total Number of Observations by Variable and Percent by Demographic

Variable	N	% FRL	% ELL	% IEP	% Minority	% Female
Total	9990	60%	57%	6%	77%	56%
ADA	9990	60%	57%	6%	77%	56%
AP Any	1193	38%	35%	0%	54%	51%
AP EBRW	269	39%	35%	0%	65%	56%
AP Math	228	31%	35%	0%	40%	36%
CE Any	6080	57%	56%	1%	72%	62%
CE ELA	1727	51%	47%	1%	69%	62%
CE Math	1202	58%	58%	0%	74%	61%
IB Any	751	79%	81%	0%	77%	75%
IB ELA	56	80%	73%	0%	79%	73%
IB Math	64	77%	75%	0%	80%	64%
PBA Any	1312	64%	53%	7%	82%	53%
PBA English	564	59%	45%	8%	77%	46%
PBA Math	748	68%	59%	6%	85%	58%
PSAT/SAT EBRW	4407	59%	55%	4%	76%	57%
PSAT/SAT Math	4407	59%	55%	4%	76%	57%
PSAT/SAT Total	4407	59%	55%	4%	76%	57%
Post-HS Plans	3051	57%	55%	5%	78%	56%
GPA	9990	60%	57%	6%	77%	56%