Using Current Assessments in District-Determined Measures:

Leveraging the Curriculum-Embedded Performance Assessments From the Model Curriculum Units

October 2013

Contents

Introduction

On June 28, 2011, the Massachusetts Department of Elementary and Secondary Education (ESE) adopted new regulations to guide the evaluation of all licensed educators: superintendents, principals, other administrators, teachers, and specialized instructional support personnel. In the implementation of these regulations, districts are selecting measures of student learning, growth, and achievement to determine the Student Impact Rating component of educators' evaluations. These measures are called district-determined measures (DDMs). DDMs provide

> This guide presents a process for building a DDM by bringing together existing assessments. It includes two examples using ESE's model curriculum units.

information to educators about student learning, the effectiveness of curriculum, and educator impact. Districts should select DDMs that are **aligned to content and informative.**[1] The regulations provide flexibility as related to the measures used in DDMs. This flexibility allows districts the opportunity to identify DDMs most aligned to the curriculum and informative to educators. DDMs can be new locally designed assessments, commercial assessments, or a modification of existing assessments.[2]

Including teams of educators with deep knowledge about students and curriculum in the selection of DDMs supports the priorities that DDMs are informative as well as aligned to the content. DDMs should provide useful information about student learning and curriculum effectiveness as well as provide evidence about a teacher's impact on student growth. Involving experienced educators helps ensure that DDMs provide useful information in these areas. Although parts of the process of selecting DDMs are new, educators have the necessary skills for identifying high-quality DDMs. In fact, educators use the skills outlined in the following three steps all the time in the course of identifying or developing classroom assessments. Experienced educators, even without specialized training, should play a central role in the identification of DDMs.

> **Challenges to Using Curriculum Assessments**
>
> 1. Using measures of achievement to measure growth
>
> 2. Using unit assessments to make inferences about the entire year

The focus of this guidance document is to help district staff understand how to build a DDM by bringing together existing assessments to create one measure of an educator's impact on student learning. This approach has many advantages. First, it is often more feasible for a district to use existing assessments when compared to the costs associated with commercial assessments and the time required to build new assessments. Second, choosing to modify familiar assessments allows educators to build on a valuable preexisting base of

---

[1] This is the central message of *Technical Guide B,* available at http://www.doe.mass.edu/edeval/ddm/TechnicalGuideB.pdf.

[2] We have referred to these approaches as "build, buy, or borrow." This document features a strategy for DDM development that best fits into the "borrowing" category. See Technical Guide A, pages 22–36, for a deeper discussion of approaches to identifying DDMs, available at http://www.doe.mass.edu/edeval/ddm/TechnicalGuide.pdf.

expertise. Finally, because these assessments are already in use, educators have most likely already completed the work of aligning the assessments to the curricula and determining that the assessments are sufficiently informative.

Using current assessments as a basis for a DDM presents two challenges. First, many of these assessments are measures of student achievement, not student growth. Measuring educator impact based on student achievement alone fails to take into account a student's wide range of prior academic performance. In contrast, measuring growth ensures that educators teaching low-achieving students will be just as likely as educators teaching high-achieving students to receive an impact rating of moderate or high. In fact, unlike measures of student achievement, measures of student growth actually create opportunities to identify where low-achieving students are making tremendous gains. Another challenge to using current assessments is that many assessments cover only a short period of instructional time. For example, a two-week curriculum unit with an assessment at the end of the unit would not be appropriate by itself to measure a year's worth of growth.

The examples used in this document show how to take existing end-of-unit assessments and adapt them so that they are appropriate to use as a DDM. These assessments are a great starting place, but it is important to remember that any DDM must be reflective of learning throughout the year. Because DDMs are used to make inferences about a "year's worth of student growth," bringing together multiple unit assessments throughout the year is one approach to constructing a DDM that supports this use.

Using Current Assessments in DDMs

The guidance in *Technical Guide B* outlines a process for piloting and implementing DDMs. This guide, *Using Current Assessments in District-Determined Measures,* presents specific examples to illustrate steps for building a DDM by bringing together existing assessments to create one measure. The following two examples demonstrate how this process can be used when linking together multiple existing assessments or expanding one assessment as a basis for a DDM.

This process includes the following three steps.

1. **Identifying key content:** What content best represents a meaningful sample of the learning a student should complete during the year?

2. **Ensuring that change in performance represents growth:** Are the assessments similar enough to support meaningful inferences about student growth during the year?

3. **Selecting an approach for measuring growth:** What scoring approach best captures student learning?

Identifying Key Content
Experienced educators will be able to look across the goals and standards covered in different units of instruction to identify important content. This process is similar to the work that educators do each day when planning units and lesson plans.

Ensuring That Change in Performance Represents Growth
Ensuring that assessments are similar enough to support claims of growth requires careful attention and professional judgment. For example, a student's performance on an oral report compared to a written pretest may be more a result of a student's talent in public speaking. As a

result, this change does not support a claim of growth. Teachers are often asked to make this type of professional judgment.

Selecting an Approach for Measuring Growth
The final step of this process involves choosing a method to measure growth. *Technical Guide B* introduces four practical approaches to measuring growth that may be familiar to many educators. These approaches are practical and straightforward and do not require sophisticated mathematical calculations to determine growth.

Practical Examples: Model Curriculum Units

The other purpose of this guide is to provide concrete examples of what this process would look like in practice. This guide provides two examples using the model curriculum units (MCUs).

These units include curriculum-embedded performance assessments (CEPAs) that are consistent with the shift in assessment emphasis represented in the 2011 Massachusetts Curriculum Frameworks. These assessments are well aligned to the frameworks and target a student's ability to use the information they have learned. Due to these properties, CEPAs make an excellent starting place for constructing a DDM. Districts interested in using the CEPAs and other existing assessments as a starting place for building a DDM can use the process outlined in this guide.

The hypothetical district, used in the examples that follow, is a moderate-size district with three elementary schools, one middle school, and one high school. The district does not have any personnel dedicated to assessment, budgets are limited, and the district is working to implement DDMs with minimal changes to existing practice.

In the first example, the principal of one of the elementary schools leads a team developing a DDM based on the CEPAs included in the nine MCUs in third-grade English language arts (ELA) that the district has adopted. These MCUs include units from both science and social studies that incorporate standards from ELA. This DDM will apply to seven teachers across the three different elementary schools. As a result, the team constructing the DDM is relatively large. This example demonstrates how a district with assessments from across a full year could survey multiple units to construct a DDM.

> **MCUs**
>
> MCUs were developed to promote effective instruction to improve student learning. They provide districts and teachers with high-quality and rigorous units for instruction or in development efforts. These units were
>
> - Developed by educators in more than 60 districts from across the state
> - Aligned to the 2011 Massachusetts Curriculum Frameworks in ELA, history and social studies, mathematics, and science and technology and engineering
> - Developed using the "Understanding by Design" framework
> - Designed with CEPAs to assess student learning

In the second example, two high school social studies teachers construct a DDM. This DDM assesses the same content area (writing), but it is intentionally different. The four MCUs used in this example were adapted for ninth grade from units initially designed for eighth, ninth, and tenth grade.

The purpose for providing two different examples is to demonstrate that this process could look different in each district. Although the examples can inform their practice, the reader's attention should **focus on the process and questions associated with each step.** It is important to remember that no example covers all of the potential challenges districts may face in applying the process described herein. Districts will bring different resources to this work and will have different goals. It is not the intention of this document to suggest that any curriculum, assessment, or approach is better suited for DDMs than another. **Instead, this document provides examples of the steps, thinking, resources, and questions that can guide educators in the process of identifying and selecting DDMs.** Districts can execute this work with confidence because, with a thoughtful team of experienced educators, identifying and building excellent DDMs is within the reach of every district.

Example: Third-Grade ELA

Creating the Working Group

Mary, the principal at Woodland Example Elementary, is leading the effort for her district in identifying a DDM for third-grade ELA. Before she starts, she knows that she needs to include the right people in the working group. She is working to include not only people with a certain set of skills, but to form a group that is representative of the diversity of perspectives from across the district. She invites a third-grade teacher from each of the other two schools in the district. These teachers bring deep knowledge about the curriculum. She also invites a second-grade teacher with experience using the curriculum in previous years. Next, she includes a teacher with experience in assessment. The reading specialist brings a broad knowledge of all the district

> **Working Groups Include Individuals Knowledgeable In**
>
> 1. Curriculum content
> 2. Assessment
> 3. Special populations (ELL, special education)
> 4. ESE's resources
> 5. Evaluation

assessments and rubrics to the team. Mary also includes an English language learner (ELL) specialist and a special education teacher to help ensure that the DDM is accessible and fair to all students. Finally, she joins the team herself to help guide the process and ensure that the measure provides useful information for educator evaluation.

Identifying Key Content

What content best represents a meaningful sample of the learning a student should complete during the year?

After assembling her team, Mary knows that selecting the right content to measure is the next step in identifying a DDM. A DDM that does not measure important content cannot provide useful information to educators about student learning. The selected content does not need to be a comprehensive collection of every learning goal covered in a year but instead must be a representative sample of the knowledge and skills taught by an educator. As Mary's team looks over the multiple unit plans, the working group notes the diversity of options. Each MCU includes many different types of goals and standards, many of which do not overlap across units. How

should the group narrow down all the many goals and standards to a focused list? This is the challenge many districts will face when looking across multiple units.

Mary reminds the group of the importance of approaching this work systematically. To help the team focus on what content would be appropriate for a DDM, she introduces the team to the process of developing core course objectives (CCOs)[3] Educators begin by generating 10–15 learning objectives and then work on reviewing and refining these objectives. This process helps teams balance several important criteria, including the content's importance, specificity, and ease of measuring. After learning about this process, the team decides to focus on the concept of "writing clearly" as the content for the DDM. This content represents important skills taught during the year and that teachers have experiences assessing. The group raises two concerns around this content. First, "writing clearly" is too broad an area to assess in a DDM. Second, writing clearly, by itself, does not identify specific standards within the curriculum framework. To address these issues, the group looks across the writing standards covered in each unit.

One advantage of using the MCUs is that they are designed explicitly to support alignment with the Massachusetts Curriculum Frameworks. The standards covered in each unit are identified at the beginning of each unit. The group decides to map out the writing standards covered in each unit. The content of a DDM should be a meaningful sample of important content from during the year, and the team wants to know what standards were visited during the year. To answer this question, the team generates the following table. From the table it is clear that standard three is covered by five of the units, more than any other standard. Standard three reads: "Write narratives to develop real or imagined experiences or events using effective technique, descriptive details, and clear event sequences." Although there is no guarantee that the standard with the greatest frequency is always the best choice, the team agrees that this standard represents easily assessed, important content covered during the whole year. The team agrees that focusing on how students perform relevant to this standard would provide important

MCU Name	Writing Standards Covered in Each MCU									
	1	2	**3**	4	5	6	7	8	9	10
Independent Readers					X					
Author Study		X	**X**	X						X
Biography	X		**X**	X						X
Extreme Weather		X					X			
Newspaper			**X**		X	X				
Poetry			**X**							
Reading Inquiry				X	X					
Stories Matter										
Whose Story		X	**X**							

---

[3] You can learn more about this process on ESE's website:
http://www.doe.mass.edu/edeval/ddm/CoreCourseObjectives.pdf.

instructional information about student progress and provide important information that should inform future curriculum revisions.

Ensuring That Change in Performance Represents Growth

Are the assessments similar enough to support meaningful inferences about student growth during the year?

After the team agrees on the content to assess, the next task is to look across the assessments used in each unit to ensure that the assessments are comparable enough to support a claim of growth.[4] Although assessments do not need to be identical, whenever possible similar tasks and rubrics support inferences about student growth. If assessments are very different, it is difficult to claim that changes in performance are due to student growth and not due to the differences of the tasks. On the other hand, identical tasks not aligned with content do not provide useful information. An educator's professional judgment needs to balance these competing goals. If educators are comfortable that improvement in performance across time is a reasonable reflection of student growth, then the assessments are adequately similar.

MCU Name	Rubric Items					
	1	2	3	4	5	6
Author Study	Topic Development and Clarity	Evidence and Content Accuracy	Organization	Standard English Conventions	Variety of Sentences	Precise Use of Vocabulary
Biography	Organization	Opinion	Illustrations	Text Features	Voice	
Newspaper	The Writing Process	Use of Technology	Language and Conventions	Interviewing and Quotes	Descriptive Details	Organization
Poetry	Includes Description of Topic of Poem	Includes Senses or Emotions Evoked by Poem	Specific Words or Phrases to Support Response	Personal Reflection		
Whose Story	Topic Development	Evidence and Content Accuracy	Illustrations	Standard English Conventions		

The team looks across the five CEPAs that include writing standard three. Each CEPA includes a writing task assessed by a rubric. However, each rubric was designed independently for each CEPA. As a result, the rubrics are not identical from unit to unit. Often when districts are using multiple assessments, even when assessing the same content, they will face challenges such as

---

[4] It is important to note that comparability is also a consideration for districts when looking across different DDMs. When looking at comparability across DDMs, the inference that districts are making is that a growth rating on one DDM is comparable with a growth rating from a different DDM. In this context, we are looking at the comparability of assessments within a DDM. As a result, the guiding inference is different. In this context, assessments need to be similar to support a conclusion that different levels of performance are a reflection of growth.

this. The team begins by producing a table that includes the rubric titles for each CEPA. They quickly notice that two rubric items appear in multiple units.

As the team looks into the descriptions included in each rubric item, they notice that despite differences in wording, content is assessed consistently across the different rubrics. The resulting table follows. The team finds four categories that are consistent across the five units: Organization (in red), Writing Conventions (in yellow), Details (pink), and Topic Development (green.) The team identifies rubric items in the category Details when they assess a student's ability to provide evidence to support a claim. These skills are included in the category Evidence and Content Accuracy, in the MCU Author Study, and Opinion in the Biography unit. Similarly, the category Topic Development identifies rubric items that assess a student's ability to tie the whole writing piece together into a single narrative. This content is included in categories labeled Topic Development, Voice, and Personal Reflection. Not every category is represented in each rubric, and in one case a single rubric item includes elements categorized in Organization and Topic Development. In addition, each rubric includes items that are unique to that specific unit.

Unit Name	Rubric Items					
	1	2	3	4	5	6
Author Study	Topic Development and Clarity	Evidence and Content Accuracy	Organization	Standard English Conventions	Variety of Sentences	Precise Use of Vocabulary
Biography	Organization	Opinion	Illustrations	Text Features	Voice	
Newspaper	The Writing Process	Use of Technology	Language and Conventions	Interviewing and Quotes	Descriptive Details	Organization
Poetry	Includes Description of Topic of Poem	Includes Senses or Emotions Evoked by Poem	Specific Words or Phrases to Support Response	Personal Reflection		
Whose Story	Topic Development	Evidence and Content Accuracy	Illustrations	Standard English Conventions		

The working group agrees that these four categories are important and relevant aspects of the assessment goal of communicating effectively in narrative writing. To ensure comparability, the team agrees to modify the rubrics for each CEPA. As a team, they agree on wording for these four categories. They add these four rubric items to all five CEPA units. These categories replace the similar categories identified previously when appropriate. For these common theme rubrics, they create rubric items with five levels for each category. This decision is made to provide greater opportunity for students of all ability levels to demonstrate growth.

The solution of creating consistent rubric items across each CEPA has several advantages in this specific situation. It allows for comparability without changing the assessments themselves. As a result, teachers will continue to use the lessons and CEPAs as originally designed. Changes made to the rubrics will not change a student's experience of the units. Additional rubric items that were applicable to only one unit, such as the use of technology in the MCU about Newspapers, can still be included. However, by adding rubric items with identical wording, it

allows educators to feel confident that changes in a student's level of achievement is the result of growth in performance and not the result of differences in the wording of a rubric item.

Selecting Approach for Measuring Growth

What scoring approach best captures students' learning?

The final challenge facing the working group is determining how the rubrics can be used to measure growth. *Technical Guide B* outlines four different broad approaches for measuring growth that are practical for DDMs: (1) using a pretest and a posttest, (2) using repeated measures, (3) using a holistic rubric, or (4) using a posttest only. In a pretest and posttest model, a student's growth is reflected by the difference between the scores on the two different assessments at two different points of time. Repeated measures is a similar approach, except with many different administrations, that allows teachers to gain information about how a student has performed across the whole year. Using a holistic evaluation is different from repeated measures because instead of grading each assessment individually and then looking at changes in scores, educators would use a holistic rubric across multiple original pieces of student work. Finally, a posttest-only approach uses other information about a student to allow for a fair comparison of a student's work. The working group reads and discusses detailed descriptions of these approaches to measuring student growth in *Technical Guide B.*

| **Four Approaches** |
| Pretest and Posttest |
| Repeated measures |
| Holistic evaluation |
| Posttest only |

DDM results are used to inform instruction, evaluate curriculum, and provide evidence about a teacher's impact. A high-quality DDM should provide information for all of these uses. As a result, the group arrives at a solution to measuring growth that fits their DDM. Elements of the DDM the team develops would be classified as a pretest and posttest, while another aspect of the DDM would be repeated measures. This highlights the fact that these four methods are designed to help educators think creatively about ways to measure growth. Because DDMs are used to determine an educator's Student Impact Rating, the team decides that these assessment should be graded by educators other than the teacher who administers the assessment. However, there is concern that it is not feasible to have outside raters for all five CEPAs. As a result, the first and last CEPAs administered to students would be rated by two other third-grade teachers from the district. The group discusses whether raters should examine differences in their scores. The group decides against this for two reasons. First, with so many assessments, this would be time consuming. Second, by having each rater's scores recorded separately, it will be possible to explore if there are systematic differences in the way individual raters scored. Even though the team plans to include a systematic training of raters and use anchor papers[5] as references, it is still critical to evaluate the consistency of raters.

Additional Considerations

DDMs also provide educators with important information about their students. By only grading the first and last CEPAs, a lot of important information is lost. The CEPAs during the year provide

---

[5] Anchor papers are examples of student work that demonstrate student performance at each of the different levels of a rubric.

information about how well students are progressing toward the end-of-year goal. Teachers can use this information to see if units help move students forward in this area. As a result, the team decides that the classroom teacher will still rate all five CEPAs, providing consistency across the entire year.

The team discusses that this method does not take into account all the information gathered from the other time points when determining an educator's impact. One idea is to allow a classroom teacher to add another CEPA when a student's work is not an accurate representation. In those cases, a student's score will be the lower of the scores on the first two CEPAs or the higher of the last two CEPA scores. However, this makes the overall system too complicated, and the group's final decision is to base scores just on the first and last CEPA. It is acknowledged that although not all of the information from each time point is used to measure impact, the results still provide worthwhile information to teachers.

Finally, the special education teacher raises another concern. Even though students will be granted any accommodations included in that student's IEP, she worries that it may not be as easy for special education students to grow across the CEPAs as it would be for other students. Looking across the rubrics, the working group disagrees about whether this would be the case. As a result, the team decides that question would be a focus for the pilot.[6]

Because the group has five different CEPAs evaluated with the same four rubric categories, they decide this DDM would provide excellent feedback for parents about how students are progressing throughout the year. The special education teacher, who has experience with spreadsheets, creates a template that produces graphs for each student showing how that student compared to the class as a whole. The working group decides that these graphs would be excellent to share during the midyear parent-teacher conferences.

Example: Writing to Text: Ninth-Grade U.S. History

Creating the Working Group

Tyson, the principal at Example High School, has been charged by the district with supervising the selection of a DDM for "writing to text" in social studies. He knows that the working group should include members knowledgeable about the curriculum, assessment, special populations, the purpose of DDMs, and the resources provided by the ESE. Although Tyson cannot personally attend multiple meetings, he is confident that his two U.S. history teachers are knowledgeable about curriculum and assessment. He decides to sit down with the two teachers and share the goals of DDMs along with multiple ESE resources he has collected. Tyson asks the two teachers to meet together and draft a "writing to text" DDM proposal. Once they have drafted a plan, he asks them to have the ELL specialist and the special education team review the proposal. After these reviews, Tyson will meet with the teachers to give final approval.

---

[6] Focus questions are discussed in *Technical Guide B,* and webinar 5. As a part of continuous improvement, districts should always have meaningful questions that they are trying to address with DDMs, both for the pilot and moving forward into the future.

Identifying Key Content

What content best represents a meaningful sample of the learning a student should complete during the year?

The two U.S. history teachers determine that their first task will be to identify the content they will assess. Both teachers have been paying close attention to the shifts in the Common Core State Standards and know that "writing to text" is a practice where students include evidence from reading materials. The teachers reflect that, in the past, the course in which they have focused most on these types of writing skills has been ninth-grade U.S. history. These skills are important in building a foundation for future success in high school social studies courses.

Because the teachers know that the MCU units include clear identification of the standards covered in each unit, they produce the following chart to create an overview. They use this overview to evaluate alignment to the concept of "writing to text." Because these units have been modified from units originally designed for different grade levels, the identified standards are not aligned. However, the underlying content is consistent and well represented in each of the four units. The results are included in the following table. Looking across the standards covered in each unit, the teachers notice that although the standards cited are not exact, the content covered by each of these standards is similar. They agree that the standard **HSS Writing.9-10-1** best represents the standard that they cover in each of the units after modifying the units. They agree that this is an important skill. They also agree that it would be critical that the second DDM used for the teacher's Student Impact Rating be focused on content.

MCU Name	Standards
Constitutional Rights (Ninth Grade)	**ELA9-10 W.1:** Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence.
Economies of North and South (10th Grade)	**RH 10.1:** Cite specific textual evidence to support analysis of primary and secondary sources, attending to such features as date and origin of the information. **WHST 10.9**: Draw evidence from informational texts to support analysis, reflection and research.
Industrial Revolution (Eighth Grade)	**W.H. 6-8.1 b:** Support claims with logical reasoning and relevant, accurate data and evidence that demonstrate an understanding of the topic or text, using credible sources.
Westward Expansion (U.S. History 1)	**HSS Writing.9-10-1**: Introduce precise claim(s), distinguish the claim(s) from alternate or opposing claims, and create an organization that establishes clear relationships among the claim(s), counterclaims, reasons, and evidence.

Ensuring That Change in Performance Represents Growth

Are the assessments similar enough to support meaningful inferences about student growth during the year?

After choosing the content for the DDM, the two teachers look through the CEPAs included in each of the four units. Several problems arise with using all of the CEPAs in this situation. First, the assessments in each unit are considerably different and are not always assessments of writing. For example, in the unit on Western Expansion, the CEPA includes a small-group project creating an exhibit, and in the unit on the Constitution, the CEPA asks students to participate in a presentation where they take on different roles. The teachers like both of these assessments. The assessments would get students involved and push students to think more deeply about the content. However, these CEPAs do not focus on writing to text.

One of the two teachers has a great idea. In the unit on the Industrial Revolution, students produce a written argument as if they were writing an editorial for a newspaper. The teachers decide that a version of this CEPA would support the existing CEPAs already included with each unit. For example, in the unit on Western Expansion, the CEPA asks students to choose items that they would be included in a museum exhibit. The teachers decide to create prompts that ask students to make a written argument for why the item they chose should be included in a museum exhibit. During the creation of the exhibits, students would be thinking back to the text as they choose items. In the written response, the prompts would require students to make the connections to the text explicit.

This DDM has several advantages that the teachers like. Students still complete the CEPAs as originally designed, and the modest addition of written questions supports student learning. This consistent task provides a comparable assessment that can be used to see how students grow throughout the year.

Selecting Approach for Measuring Growth

What scoring approach best captures students' learning?

Next, the teachers consider how they will use these assessment to measure growth. They review the four approaches outlined in *Technical Guide B*. Although it is a new approach for both of them, they like the idea of collecting writing samples throughout the year and looking for growth holistically across the four samples. This method has the advantage of being able to capture changes in student's ability that could be lost by rating each piece individually. The teachers develop a holistic rubric that looks at growth across the four writing samples in four aspects of citing textual evidence. The areas are proper citation, specificity, connections between details, and relevance.[7] They decide that students will be rated

> **Four Approaches**
>
> Pretest and Posttest
>
> Repeated measures
>
> Holistic evaluation
>
> Posttest only

---

[7] Improvement in *citation* includes greater accuracy in using citation conventions. Improvement in *specificity* includes more accurate and fuller citation of sources. For example, a student who previously wrote: "Benjamin Franklin said we cannot give up liberty in order to be safe," would demonstrate improvement in specificity if later the student wrote: "In the *Memoirs of the Life and Writings of Benjamin Franklin*, published in 1818, he writes that 'they who give up essential liberty to obtain a little temporary safety deserve neither

on a three-point scale in each of the four areas. They acknowledge that not every student will grow in every area.

Because a DDM is a part of a teacher's impact rating, ideally assessments would be evaluated with external raters. However, as is the case in this situation, additional raters are not always feasible. As a result, the two teachers agree to rate all students (including each other's) independently at the end of the year. They will collect the students' writing samples throughout the year and propose that they use a professional development day at the end of the year to grade the assessments. Students who are scored differently by each teacher would be noted, and the teachers would discuss these differences. The goal will be to agree on a final rating. In their meeting with the principal, both teachers acknowledge that this method of scoring has the potential chance of bias because a teacher would likely recognize his or her own students. The two teachers agree to work in the future to continue to improve the rubric to make it as objective as possible.

The ELL specialist is concerned that the three-point scale may be problematic because it may not differentiate enough between levels of student growth. She is concerned that this may result in an unfair opportunity for students to demonstrate growth if they are focused heavily on one area. However, the two teachers are worried that they will not be able to make these types of distinctions in a consistent way. The principal and the history teachers agree that it would be important to revisit the number of points in each scale next year after they have had more experience with the DDM. Despite this concern, both the special education team and the ELL specialist like the approach in general because it takes into consideration where a student starts.

Conclusion

In each of these examples, district teams of teachers determined the measure used in a DDM. There were important similarities between these two examples. Both DDMs assess student writing and are based on the CEPAs from the MCUs. However, despite these similarities, the differences between the resulting DDMs are considerable. Resources, purposes, and what individuals contribute differ in each example. For example, in elementary classrooms, teachers have more opportunities to meet with parents, and as a result, the third-grade example provides more in-depth information for parents. On the other hand, in the high school example, the DDM provides a structured opportunity for the two teachers, who often do not have time to collaborate, to look systematically at the growth of their students. The DDM also provides an authentic opportunity to discuss areas of relative weakness and how each teacher is approaching this important area of student learning when they work together during the professional development day.

What is important about these two examples is that they provide two realistic examples of how the process of selecting DDMs would look in practice. The careful reader will notice that even in these examples, important questions arise around the bias of raters and a reasonable concern that teachers might begin to alter the curriculum to have greater focus on the content assessed in a DDM. These are important concerns of which districts need to be aware. The more districts can leverage DDMs to provide meaningful information about students and learning, the greater the chance that DDMs can be used as a tool to support student learning.

---

liberty nor safety.'" Improvement in *connections between details* means that students draw more comparisons and contrasts between details. Finally, improvement in *relevance* means that there is improvement in how a student utilizes details to support their overall argument.

It is important to note that in both of these examples the content assessed by the DDMs is the skill of writing. Skills, typically taught consistently throughout the year, make a logical choice for the content of a DDM. However, it is not the intention of this document to suggest that DDMs cannot also be measures of content. One of the challenges to using content as a DDM is that typically different units cover different content. This makes it difficult to ensure enough similarity of the assessments across the year as described in this document. One approach that would address these issues would be to create a DDM that was an average of multiple assessments of growth for each unit. These assessments could be performance assessments or paper and pencil tests. Thus for each individual pretest and posttest, a district would look at the key content, assessments that support a claim of growth, and a method of measuring growth for each unit.

DDMs are an excellent opportunity for districts to widen the focus of student learning to content not typically assessed. The process of selecting measures asks teachers to think carefully about the important question, "What should my students leave my classroom being able to do that they could not do when they walked in?" As these two examples demonstrate, DDMs will look different in different situations. Implementing DDMs requires districts to balance multiple competing goals including alignment with the curriculum, measuring quality, and overall feasibility. This can be a complex process but one that is within the capability of experienced educators.

Acknowledgements