

Common Formative Assessment

**A TOOLKIT FOR PROFESSIONAL
LEARNING COMMUNITIES AT WORK™**

Kim Bailey

Chris Jakicic

Foreword by Richard and Rebecca Doolittle



CHAPTER 2

Setting the Stage for Common Formative Assessments

KEY POINTS

- There is compelling research that says that frequent formative assessments improve student achievement for all students.
- Common formative assessments do not have to be tests or quizzes.
- Common formative assessments do not have to take a long time to administer or include lengthy student work products.
- If you don't use the results of the common formative assessment to make a difference in student learning, the assessment is summative.

As we discussed in chapter 1, four critical questions guide the work of teams in PLCs (DuFour et al., 2010a, p. 28):

1. What knowledge and skills should every student acquire as a result of this unit of instruction?
2. How will we know when each student has acquired the essential knowledge and skills?
3. How will we respond when some students do not learn?
4. How will we extend and enrich the learning for students who are already proficient?

This book will help your team confidently answer the second question by using common formative assessments. Specifically, this chapter will help define what common formative assessments are and how they fit into a well-developed, balanced assessment system.

When your team begins to create assessments to determine whether or not students have learned the identified essential learning outcomes, you are beginning the work that many believe is pivotal to the process—the work that will really make a difference for your students, but that will also challenge you as a teacher. Before teachers really understand this work, we often hear them say, “We are already testing kids too much. Why would we want to do more testing?” and “I already know which of my kids

need help. I don't need another test to tell me that." These teachers aren't being difficult; they are just expressing their concern that any instructional time they take away from the teaching process will have a negative impact on their students.

Once teachers begin this work and start to see success, they understand that writing and using common formative assessments is not *one more thing* but rather an integral part of the teaching and learning process. Let's examine what teams need to know to help them see the value of this important step in the PLC process.

Formative Versus Summative Assessment

Most teachers are comfortable that they know the difference between formative and summative assessments. They know that formative assessments are assessments *for* learning and summative assessments are assessments *of* learning (Stiggins, Arter, Chappuis, & Chappuis, 2004).

Other assessment experts have written about the difference. For example, according to Reeves (2009):

It is absolutely vital that we understand the true meaning of formative assessments—an activity designed to give meaningful feedback to students and teachers and to improve professional practice and student achievement. Tests designed only to render an evaluation cannot achieve the potential of assessment for learning that assessment experts have suggested is an essential element of effective practice. (p. 91)

DuFour, Eaker, and Karhanek (2010) clarify formative assessments even further:

Three things must occur for the assessment to be formative: (1) the assessment is used to identify students who are experiencing difficulty, (2) those students are provided additional time and support to acquire the intended skill or concept, and (3) the students are given another opportunity to demonstrate that they've learned. (p. 63)

We address this clarification in more detail throughout this book. Chapter 7 will help teams see how to use the results from these assessments to identify not only *which* students need help but also *what kind of help* they need. In that chapter, we also explore ways that teams can find the time to provide help for identified students. Later in the chapter, we also explore the idea of how teams will need to think differently about their grading practices once they begin using formative assessments.

We believe the purpose of the assessment and how teams use the results is what really determines whether it is formative or summative, not how it's written or administered. If the assessment occurs during the learning process, and the results will be used to help students continue to learn, it is considered formative. As DuFour et al. (2010b) note, *formative assessment* is "used to advance and not merely monitor each student's learning; the assessment informs the teacher regarding the effectiveness of instruction and the individual student regarding progress in becoming proficient" (p. 3). If the assessment occurs after the learning is complete, and is used to give a grade or provide a final measure of student results, it is *summative*. So the biggest difference will not be in what the assessment looks like but rather in how teachers respond to the results. For example, if an English teacher asks her students to complete a graphic organizer comparing themes of two stories, grades the assignment, and then returns them to her students believing it's a formative assessment, she's confused the difference between formative and summative. What would make this assessment formative is if she used the assignment to determine which of her students were not able to compare the themes of the stories and then provided them with additional instruction as a result of the information.

You will learn in chapters 4 and 5 that in order to develop truly effective formative assessments, you will need to break *standards*—the narrowest item listed by a state when describing what students should know and be able to do—into each of the learning targets that are made clear to students. *Learning targets* are the smaller skills, strategies, and pieces of content information a student needs to know in order to be able to complete the standard (see figure 2.1). The process your team will use to carefully uncover these learning targets is described in chapter 4, *The Unwrapping Process: Achieving Collective Clarity on Learning Targets*.

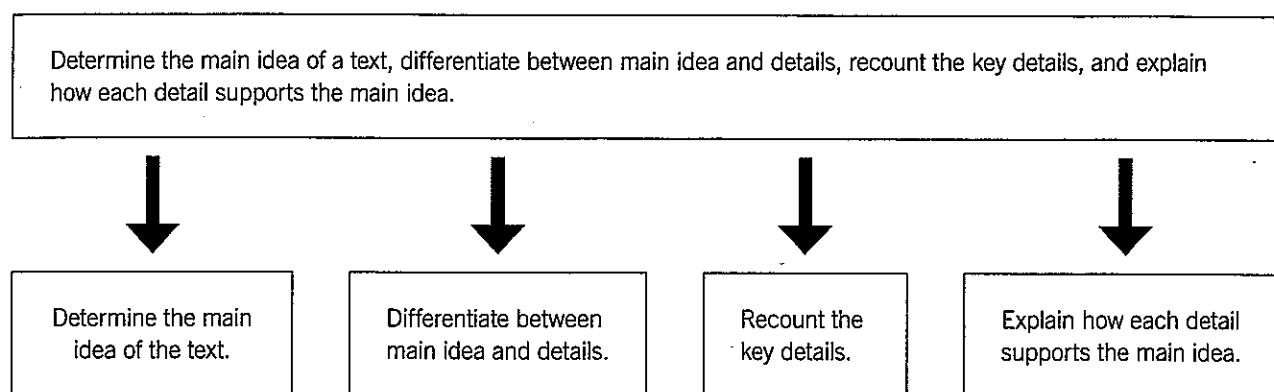


Figure 2.1: Example of how a learning standard breaks down into learning targets.

The term *standard* is used throughout this chapter in a generic way. The Common Core State Standards Initiative uses the term *standard* to "define the knowledge and skills students should have within their K–12 education careers so that they will graduate high school able to succeed in entry-level, credit-bearing academic college courses and in workforce training programs" (Common Core State Standards Initiative, 2010a). However, before these standards were developed, most states used a variety of terminology to mean this same thing. For example, Arizona uses the term *performance objective*, and Illinois uses the term *performance descriptor*. And to make it even more complicated for teachers, even within the same state, standards often are written with a different *grain size*. That is, one might be a specific skill, while another, a much larger learning outcome.

This concept of unwrapping is one of the key strategies teams can use to write formative assessments to guide their instruction. A formative assessment is usually written around learning targets (specific skills and strategies), and a summative assessment is usually written around more complex standards or even multiple standards. The reason that this is important is that the learning targets are the step-by-step processes we teach students as they move toward their understanding of a bigger concept. We assess them formatively so that we know exactly how to respond when a student is experiencing difficulty during the learning process.

In chapter 5, your team will learn a process for designing a formative assessment around a small number of learning targets so that you can identify and respond to students who haven't learned the targets in an expeditious way.

Some common misconceptions many teams have about formative assessments are that they have to be written, they have to be tests or quizzes, or they have to be lengthy. Formative assessments can be student work samples and activities previously used by teachers as practice during the unit of instruction. They can also be performances teachers watch and score against a rubric. Effective formative assessments are intentionally short in terms of the number of questions or items, and they should take only a short time for students to complete. Because we want teams to feel like their assessment practices are a seamless part of teaching and learning, and because we want them to occur frequently, it is important that they be short enough in length that they don't interrupt the instructional process. We encourage teachers to design assessments that take as few as fifteen to twenty minutes to administer.

What Does *Common* Mean?

The term *common assessment* refers to those assessments given by teacher teams who teach the same content or grade level—those with “collective responsibility for the learning of a group of students who are expected to acquire the same knowledge and skills” (DuFour et al., 2010b, p. 2). For the assessment to be common, no teacher can opt out of the process; it must be common to all teachers who teach that course or grade level. DuFour et al. (2010b) add that common assessments use “the same instrument or a common process utilizing the same criteria for determining the quality of student work” (p. 2). This means that if a rubric is used, the teachers must work on building not only a common understanding of what the rubric means but also use collaborative scoring practices to ensure the results are the same no matter which teacher applied the rubric.

Benefits of Common Assessments

The benefits of common formative assessments are great. DuFour et al. (2010a, p. 80) state that common formative assessments do the following:

- Promote efficiency for teachers
- Promote equity for students
- Provide an effective strategy for determining whether the guaranteed curriculum is being taught and, more importantly, learned
- Inform the practice of individual teachers
- Build a team's capacity to improve its program
- Facilitate a systematic, collective response to student who are experiencing difficulty
- Offer the most powerful tool for changing adult behavior and practice

Some teachers are concerned that collaborating to do this work will take away from the time they have to prepare for their classrooms. However, when considering the amount of time each teacher currently puts into the assessment process, doing the work collaboratively should result in more efficient work. Rather than starting from scratch, many teams begin the process by sharing their current assessments and choosing those ideas and items that appear to be most effective. Your team will likely experience a shift in how you work together as you move from *sharing* your current practice to actually building assessments together. Teams that design common formative assessments have more in-depth discussions about proficiency and, as they analyze the resulting data, have more focused conversations about instructional strategies (Graham & Ferriter, 2008).

One of the principles of PLCs is that teams engage in collective inquiry—that they learn together. They build shared knowledge around best practices so that they increase student learning (DuFour et al., 2008). Consider how your team will learn together as you write and use formative assessments. You will likely become better assessment designers and will strengthen your skills in analyzing and using data.

While analyzing data, your team will share instructional strategies and determine if some of those strategies are more effective than others. Even if there is no one strategy that works best, you will have an opportunity to add to your repertoire of strategies so that you can respond in a different way to the students who didn't learn it the first time.

When your team works to answer question one—What knowledge and skills should every student acquire as a result of this unit of instruction?—you begin the process of assuring equity for all of the students your team serves. In chapter 3, we'll lay out a process that teams use to build consensus about how to answer this first important question. This guaranteed and viable curriculum (Marzano, 2003) becomes the focus for your formative assessments.

Equity means that your students will learn the same important learning targets no matter which teacher they have. While this concept sounds good for students, it also has an important benefit for teachers. That is, when students are commonly prepared for the next course or grade level, as happens when they get equitable content, the next teacher doesn't have to use valuable instructional time filling in content for some students so that all students have the same prerequisite skills and information. Common formative assessments help ensure equity because they are written around agreed-on learning targets. One of the important steps your team will take while developing these assessments is to discuss what proficiency will look like. Coming to consensus about proficiency is critical for guaranteeing equity for your students.

In addition, your team will have a greater capacity to respond when students need more time and support. DuFour et al. (2010a) argue that when students experience difficulty learning, members of collaborative teams must provide additional time and support through an intervention system that is timely, directive, and systematic. They describe many examples of systematic interventions that effectively meet these characteristics. Interventions may range from differentiated instructional strategies, smaller grouping to monitor incomplete work, and even enrichment opportunities, to name a few. For example, after analyzing the results of their common formative assessments, teams may move students from one classroom into another *for a short time* to provide specific and targeted learning support. If there are three classes on the team, each teacher takes one group of students for response. One teacher takes all of the students who need more time and support with an opportunity for teaching the target using an alternate instructional strategy, another takes the students who need additional practice, and the third teacher takes those who could benefit from enrichment.

When talking about intervention, we use the term *corrective instruction* to describe strategies teams can use to respond to the results of their common formative assessments. *Corrective instruction* is instruction that occurs for some students whose assessment results indicate they have not learned a particular skill or strategy at a proficient level. It is instruction that is *different* than the initial instruction in a way that teachers believe will help the student understand and learn the skill or strategy.

The response to intervention (RTI) model is built on the premise that not all students learn the same way and at the same speed. PLCs build a system of responses to assure students will be provided the time and the instructional strategies they need to be successful. They embrace formative assessment as a way to identify which students need additional support and which students need extension or enrichment. Therefore, after initial instruction, the team administers a common formative assessment to identify which students require corrective instruction. These students are then provided with the necessary help. For some students, this initial extra support still won't be enough. Teams monitor students by using additional progress monitoring to identify which students need even more support.

Finally, your team will benefit from the professional learning that takes place as you talk in more detail after the assessment about effective instructional strategies, effective assessment strategies, and effective ways to respond to students.

As we have described, common assessments can be either formative or summative depending on their purpose. A common summative assessment, for example, might be a final exam that the entire biology team uses at the end of the year in high school. Remember that these are considered summative *because* they occur at the end of the learning and their purpose is *merely* to give a grade. In chapter 7, we'll explore ways that teams could use the results of their tests *at the end of each unit* in a formative way to provide even more time and support for the students who still haven't learned it after initial instruction.

Putting It Together

PLCs focus on common formative assessments for their work together because of the compelling research that these are the assessments that can truly improve student achievement.

In their meta-analysis on classroom assessment, Black and Wiliam (1998) report a 0.4 to 0.7 standard deviation increase in student achievement with the use of frequent formative assessment. This research was conducted for all grade and age levels. The most improvement was seen for the lowest-achieving students. Wiliam (2007) follows up by saying, "When implemented well, formative assessment can effectively double the speed of student learning" (pp. 36–37).

Popham (2006) explains it is important that this research not be misinterpreted to mean that any kind of testing will provide similar results:

Educators need to realize that the research rationale for formative assessments is based on short cycle assessments. Such rapid-turnaround assessments yield results during a class period or in the midst of a multiweek instructional unit. If the results don't get back in time for teachers to adjust instruction for the students being assessed, then it's not formative assessment. (p. 86)

Finally, as Wiggins (2006) so eloquently reminds us: "The more you teach without finding out who understands the information and who doesn't, the greater the likelihood that only already-proficient students will succeed." If we *don't* use formative assessment, we are basically regressing back to the "I taught it, they just didn't learn it" mentality. And in the end, the students who need it the most will lose out.

A Balanced Assessment System

Because of the strong research base for using formative assessments, schools are successfully creating these team-based assessments in addition to a variety of other types of assessments. They want to have all the necessary information they need about their students. Therefore, they recognize the importance of a balanced assessment system.

In a balanced assessment system, teachers have access to both formative and summative information in order to make short- and long-term decisions to help their students. So what does a balanced assessment system look like, and why do we need each type of assessment?

Table 2.1 shows the pieces of a balanced assessment system. They include classroom assessments, common formative assessments, interim or benchmark assessments, and external summative assessments. (This Balanced Assessment System Framework also appears on page 101 of the Tools for Teams appendix.)

Table 2.1: A Balanced Assessment System

	Classroom Assessments		Common Formative Assessments	Interim or Benchmark Assessments	External Summative Assessments
Examples of practice	Worksheets, clickers, whiteboards, exit slips, conferences	Final exams, final projects	Tasks assessed with rubrics, short quizzes, common worksheets, and clickers	Quarterly tests or performances, writing samples	State tests, and ACT, SAT, and Advanced Placement (AP) exams
Formative or summative?	Very formative	More summative	Very formative	More summative	Summative
Whose responsibility?	Classroom teachers	Classroom teachers	Collaborative teams at each school	District teams of representative teachers	An external group of "experts"
Purpose?	To give immediate feedback	To give a grade	To determine if students have learned the material and how to responding	To assess curriculum, instructional strategies, and pacing	To determine whether curriculum, instructional strategies, and pacing were appropriate

Classroom Assessments

First, there is nothing about common formative assessment that precludes teachers from using their own classroom assessments in addition to common team-developed assessments. In fact, we encourage teachers to use a variety of strategies to gather immediate feedback during each lesson about whether students understand the concepts being taught. Good instructional practices that provide immediate information include checking for understanding with questions; monitoring students as they work independently and in groups to provide feedback and support; using clickers (each student has a device connected to a hardware system that he or she uses to respond to questions so the teacher can see who responded, in what way, and how many students were correct or incorrect), whiteboards, and exit slips (or "tickets out the door") that ask a few questions about the lesson; as well as administering teacher-created

quizzes. Whenever teachers engage their students in practices that provide the teacher (as well as students) with information about whether they have successfully learned a target, the practice is considered an assessment. For example, if teachers ask their students to respond to the question, "What is the difference between a plant and animal cell?" as an exit slip, they can easily sort the responses into groups of who did or did not learn the difference. Because the teacher knows *by student* who needs more time, the practice is considered a formative assessment. Thus, classroom formative assessments are still necessary in the work of PLCs.

Summative assessments happen at the end of the learning process and are used to give grades. In addition, because formative assessments are often written around more isolated learning targets, summative assessments are intended to make sure that students can put skills and strategies together to be able to accomplish what we typically call the *standards of learning*. For example, a science department may have an overall goal that students can design and carry out an experiment following the scientific method to gather good information to prove a hypothesis. Teachers may create a number of formative assessments around specific learning targets: Can the student create a hypothesis? Does the student know a number of different ways to gather and arrange data? Can the student use correct safety procedures in the lab? The summative assessment, however, requires the student to be able to put all of these skills together to answer a specific scientific question.

Consider, for example, the following standard from the sixth- through eighth-grade writing standards for literacy in history/social studies, science, and technical subjects 6–12 from the Common Core State Standards (Common Core State Standards Initiative, 2010b):

Conduct short research projects to answer a question (including a self-generated question), drawing on several sources and generating additional related, focused questions that allow for multiple avenues of exploration. (p. 66)

The team might identify the following skills as important skills for students to know and be able to do in order to accomplish this standard: locate multiple sources of information about a specific topic, develop research questions to ask about a topic, and even evaluate the adequacy of the amount and type of information gathered about the topic to determine if the information is sufficient. The team could create formative assessments around one or more of these specific learning targets. However, it's possible that a student could carry out each of these learning targets and still not be able to conduct a quality research project as required in the standard. Thus, good assessment practice requires both short formative assessments around specific learning targets and final summative assessment around the learning standard.

We believe that until a student has mastered the essential outcomes being taught in a unit of instruction, there must be additional support beyond the core instruction. The core instruction includes initial classroom teaching, formative assessment around identified learning targets, and intervention—corrective instruction or extensions—around those learning targets. When the summative assessment identifies students who are not yet proficient, then the student is provided Tier 2 support for further learning. In this case, the student might be pulled out for additional small group instruction during a time where no direct instruction is occurring in the classroom. For example, a student who is still experiencing difficulty writing a coherent paragraph after the class is finished with that unit is provided specific targeted intervention on that skill. In another example, a high school algebra student who is still having difficulty understanding how to solve a quadratic equation after the unit is completed would get help in a resource class during her study hall until she demonstrates that she understands this concept. This is not to say

that instruction waits. The students who need additional support continue moving forward with the next unit along with their classmates. However, concurrently, they are receiving support in the crucial skills and concepts that are essential.

Common Formative Assessments

Common formative assessments, the assessments your team will develop and use to assure students are learning and to know what to do next when they need additional time and support, are the focus of your work as a team and the purpose of this book.

Common formative assessments are written by teacher teams around the learning targets the team has identified as the most important ones to be taught. They help your team know not only which students have learned the targets but also what to do for the students who have not. They do not have to be pencil-and-paper tests or quizzes; they can be individual student work samples, completed graphic organizers, writing pieces, products, or performances. They are written around a small number of learning targets, and therefore are not intended to take a long time to administer. Ideally, they can be scored quickly so that the team can respond in a timely fashion when it identifies students experiencing difficulty.

Consider, for example, the seventh-grade English team that is working on the following standard from the Common Core State Standards: "Support claim(s) with logical reasoning and relevant evidence, using accurate, credible sources and demonstrating an understanding of the topic or text" (Common Core State Standards Initiative, 2010b, p. 42). First, team members identify the student-friendly learning target: "I can support my position by providing evidence." Then they give their students the statement "Video games are harmful to adolescents" and ask them to take one side of the issue and find evidence to support their position. Their short formative assessment is for students to write one paragraph laying out their position and backing it up with support.

One important characteristic of common formative assessments is that any constructed-response questions or performance targets are assessed using a common rubric developed for that assessment. The team of teachers agrees on the rubric and practices collaborative scoring so that all of the team members are using the rubric in the same way. Throughout this book, we'll discuss how to design and use rubrics to provide feedback to students.

Benchmark Assessments

Many schools and districts are creating *benchmark assessments* given periodically to determine whether their students are making progress toward the mastery of standards.

It's important to note that some benchmarks are designed to monitor student attainment of specific standards that have been addressed through classroom instruction prior to its administration. Often, they are designed in alignment to a specific pacing guide. For example, the district may decide that all schools will teach five standards during the first quarter in a particular subject area. The district then creates a summative assessment to be given at the end of that quarter. Other benchmark assessments are designed more globally to monitor student growth along a recursive set of standards, some of which may not have been addressed during instruction. If your school or district is using benchmark assessments, it's critical to understand how they are structured and the purpose for their implementation.

Larry Ainsworth (2006) suggests that benchmark assessments have two benefits. The first is that they have “predictive value” for how students will do on the next level of assessment (the state test). They also help the team plan for future instruction and assessment. If the benchmark assessment is written in the same style as the state assessment or other external summative assessment (like an AP exam), students have an opportunity, before the end of the year, to practice the type of assessment item as well as know where they are on the path toward mastering each of the standards.

External Summative Assessments

Almost all of us are required to administer state assessments toward the end of the year. The purpose of these tests is to determine what percentage of our students fell into each of several levels: below proficiency, at proficiency, and exceeds proficiency. Teachers generally view these as high-stakes tests because the results are reported and publicized, often comparing school to school within a state or city. These tests are used to determine whether a school is making AYP (adequate yearly progress) toward all students being proficient. Schools that are not making AYP are often given sanctions or penalties. Other external summative assessments many schools use include AP exams, the SAT, the ACT, and achievement tests.

In a PLC, teachers believe that it is their responsibility to ensure that all kids learn. State assessments provide a common understanding of what that learning must look like for all kids. Schools use the results of this yearly assessment to determine their current reality, identify their greatest area of need (O'Neill & Conzemius, 2006), and set SMART goals for the year.

Finding the Balance

Each of these assessments is necessary for teams to create a balanced view of student learning. Teachers want to know, as they are teaching a concept, whether or not students are understanding, and they use classroom formative assessment strategies to gather that kind of feedback. Team-developed common formative assessments help teams of teachers respond—*during* the learning process—to students who need more time and support as well as those who could benefit from enrichment strategies. The team capitalizes on the benefits of collaboration to use the best strategies for instruction and corrective instruction. Periodic benchmark assessments provide checks during the school year to assure that students are making progress on proficiency toward the standards to be assessed at the end of the year on high-stakes tests. External summative assessments help schools and districts remain accountable for all students in their schools.

As teams begin the process of developing common formative assessments, they should look at their list of current assessments, considering whether they have appropriate assessments in each category and whether they have some current assessments that overlap in purpose. If so, any redundant assessments could be discontinued to provide more time for teams to work on common formative assessments. Sometimes teams will discover that they have a plethora of summative and external assessments already in place, but don't have adequate formative assessments. In this case, they should discuss the possibility of eliminating some of their summative assessments.

In many districts, there are required assessments teachers must use. Early in the development of common formative assessments, teams may find that they are doing duplicate work because their districts still require them to administer everything on the original assessment list. If this is the case for your team, you may want to keep a list of all the assessments administered over a quarter of the school

year and note the amount of student contact time it takes to administer them. This data can help you understand whether or not there really is too much time being given to assessment and whether there really is too much overlap of information. Our experience has been that sometimes duplicate summative assessments can be eliminated once administrators, teachers, and parents are confident that the common formative assessments are more beneficial, particularly when the summative assessments provide redundant information.

Do We Have Time to Do It All?

Your team may be worried about the amount of time it will take to create, administer, analyze, and respond to common formative assessments—most teams are. But many teams are already successfully doing this work and feel confident that the time they are spending is worth it. So how do they make it work?

First, we believe that teams need to have time during the contractual school day to do the work of PLCs (DuFour et al., 2010a). Once teachers learn about the compelling research behind formative assessments, they are invested in the process. The PLC model supports teacher learning, and teams that write and administer these assessments are doing some of the most powerful professional development they can do. Much of the work your team will engage in is like the work each individual teacher has done in the past. The difference is that, by doing it together, teachers benefit from the insights of the group instead of relying on only their own knowledge.

As teams get started in the process, they generally find that the steps might initially take longer than they will once the team becomes more assessment literate and has experienced a full cycle of design, implementation, analysis, and action. It's good to recognize, however, that teams that use common formative assessments on a regular basis become so familiar with the process that they can accomplish much more and their product is much better than when they first started.

For example, when starting out, teams may need a full meeting to discuss which targets of learning they will assess and when they will administer the assessment. They will likely use another meeting to write the actual assessment and discuss proficiency. Once they've administered the assessment, it will probably take an additional meeting to analyze the data and develop the plan for responding to student needs. Teams that work on this over a period of several years find that they are able to expedite the process, focusing more on the revision of assessments instead of developing new assessments. In chapter 1, we talked about how teams create a work flow from meeting to meeting. Using one meeting to write common formative assessments and the next meeting to analyze the results and plan the response is an example of how this work flow looks.

Grading Practices

When your team begins to develop and use common formative assessments, it won't be long before someone will ask, "If these assessments are being given *during the learning process*, should they be graded?" The easy answer is that they should not be graded. As O'Connor (2007) states:

Grades are broken if scores for everything students do find their way into report card grades. The fix is to include, in all but specific, limited cases, only evidence from summative

assessments intended to document learning, that is, designed to serve as assessments of learning. (p. 95)

Teachers might be concerned that if they don't grade an assignment or assessment, students won't be motivated to do their best work. However, the research clearly suggests that this is not the case. Actually, research has shown that when a grade is given, the student does *not* learn how to do better from that evaluation. In a study conducted about the motivation value of grades and feedback, one group of students was given only grades for a task, one group comments only, and one group grades plus comments. Researchers then measured students' performance on subsequent tasks. Students who were given only comments had significantly greater improvement than those given only grades or those who received both grades and comments (Butler, 1988). Corrective feedback seems to hold the power for improved learning.

Another concern some teachers have about not grading common formative assessments is that they won't have scores in their gradebook to effectively grade students at the end of the grading period. When schools use a standards-based report card, teachers can easily collect common formative assessments as evidence of the body of student work they are using to determine where the student is currently in the learning process. However, teachers in districts with a more traditional reporting system will find it hard to know what to do with the scores on these assessments. O'Connor (2007) suggests that teachers keep two parts to their gradebook—one with formative results and another with summative results. The final grade is computed using only summative scores, but the formative scores allow students and their parents to know where the students' learning is throughout the quarter.

Rick DuFour advocates a process that considers both of these teacher concerns. He suggests that collaborative teams grade the initial assessment, identifying the students who need additional time and support for learning. Then the team or school must provide the specific intervention students need to become proficient. Once a student has shown proficiency, the grade on the initial assessment is changed to reflect the new learning (DuFour, 2010).

Changing the grading system is a significant endeavor in any school or district. O'Connor (2007) calls grades "the last frontier" (p. 127). We agree with this statement and recommend that schools wait to discuss changing the grading system until the need for change emerges naturally. In other words, after teams have been working with common formative assessments for a period of time, they will have evidence of why a traditional grading system doesn't really accurately tell parents what students have learned. A revised grading system serves no purpose until you have a more accurate way to collect evidence of what students actually have learned. Teams will learn more from writing and using common formative assessments than from planning to change the grading system.

Gathering and Collecting Data

As teams increase their use of common formative assessments, they may find it beneficial to explore the idea of how to use technology to gather, sort, and store this information. We have worked with teams that use spreadsheet programs to gather and sort their data. They meet in a room with a projector and sort the data for each learning target as they discuss the results and next steps. Other schools and districts have purchased technology specifically for this use. One school district, for example, has a scanner and printer at each school so that teachers can run student answer sheets through the scanner and get a printout of the results by student and learning target. The software program aggregates (adds up) all the

results for individual teachers and prints out the data totals (see chapter 7 for what data to look for) that teachers can use to make decisions about next steps.

Before investing in expensive equipment and software, however, teams should develop and use the results of a variety of common formative assessments to determine if a software program will help them. Depending on the frequency of assessment, the types of assessment, and the number of teachers and students involved, some teams prefer to rely on their own tools to gather data and make decisions. These tools might include data tables they personally design or even Excel spreadsheets that are used to sort data they enter from each assessment. In chapter 3, we help teams begin the work of developing common formative assessments by helping them see how to answer the first critical question, What is it we want students to know and be able to do? This work becomes the foundation for the rest of the assessment process.



CHAPTER 5

Designing Quality Common Formative Assessments

KEY POINTS

- Assessments must provide information about important learning targets that are clear to students and teacher teams.
- Assessments must provide timely information for both students and teacher teams.
- Assessments must provide information that tells students and teacher teams what to do next.

After reading and using the strategies in chapter 4, you and your team know how to achieve a better understanding of your state standards and, through the unwrapping process, how to uncover the learning targets that are most important for your students to learn. Uncovering these targets is the important first step to designing quality assessments because the targets to be assessed must be clear to both the student and the teachers. This chapter focuses on the next critical step in the process: designing assessments, including writing or choosing good items, knowing what to include in the assessment, and making sure the data you collect will really help your team know what to do next to help your students.

Have you ever felt that you don't have the knowledge to design quality assessments that will effectively tell you whether or not your students have learned the skills and concepts you've been teaching? Most teachers have felt this way, perhaps because they are aware of the level of statistical analysis large test-writing companies perform, and they wonder how their skills can compare. However, assessments designed by classroom teachers that are intended to reveal what students have learned and what they need to know next around specific learning targets don't have to use the same level of statistical analysis as high-stakes tests used to determine AYP, graduation, or college entrance. Thus, for classroom teachers, it's not necessary to use what Doug Reeves (2007) calls "psychometric perfection" (p. 235). In fact, he suggests that the benefit of getting information quickly and easily by using frequent formative assessments outweighs the risks of not writing perfect assessments. He believes that because these formative assessments are done so frequently, unintended errors will be caught quickly. That is, if a student is identified as needing more help and actually understands the target of learning, the student can quickly demonstrate

learning during the corrective instruction time. Also, if a student looks like he is proficient on a target of learning, but actually isn't proficient, teachers will catch that error as they do more assessment. Of course, our goal is to make sure our assessments are as accurate as possible so that we gather good information from the beginning.

Most teams want to know how to develop practical, easy-to-use assessments that provide information about student learning and help determine next steps for classroom instruction—assessments that are valid and reliable, but that don't rely on difficult design strategies and statistical analysis.

Therefore, let's consider what the terms *valid* and *reliable* mean for common assessments. *Valid* implies that the assessment is truly measuring what the team thinks the students have learned. For example, if your team has been working toward students learning the science concept "permeability across a membrane," the assessment must tell your team which students know and understand that concept and which students haven't yet learned it. Equally as important, the assessment must measure student learning at the thinking level the team has set for student mastery. That is, your team may want students to know the definition of *permeability*, as well as to be able to describe what it means—what it looks like at a cellular level. The assessment must be designed to measure both of these thinking levels.

Teams also want to know that their assessments are *reliable*—that the students who appear to have learned the concept have actually learned it, and that the students who appear to not have mastered it truly haven't. They want to know that the results they are getting are accurate, so that they can make decisions about what to do next that will improve student learning.

Most importantly, your team wants to design quality assessments that *don't* require statistical analysis after they are administered to assure they are both valid and reliable. Rather, it wants to use a process that is set up from the beginning to assure accurate results.

There are some specific strategies your team can use for designing common formative assessments to assure you are getting the right information about what you have taught—information you can use to know what to do next for your students. This chapter presents a five-step process for designing quality assessments.

Step One: Decide What to Assess

In chapter 4, you learned how to unwrap state standards to get a better understanding of their meaning. As part of that process, you uncovered specific learning targets that your team discussed in depth, agreeing on which were most important to teach, as well as the expected thinking level for students for each target. The targets you identify as the most important become the basis for your common formative assessment plan. Formative assessments are written around these specific learning targets, rather than around state standards. However, your team does not have to assess each learning target it has identified—only those determined to be critical for students to know and be able to do. This means the team will view some of the unwrapped targets as either less important or as scaffolding skills that lead to the understanding of another target. The team decides which targets have the most value and designs the assessment around those targets.

Let's look at an example. At Emerson Elementary, the fifth-grade team is teaching the following Common Core State Standard in math:

Add and subtract fractions with unlike denominators (including mixed numbers) by replacing given fractions with equivalent fractions in such a way as to produce an equivalent sum or difference of fractions with like denominators. (Common Core State Standards Initiative, 2010c, p. 36)

Team members know there are likely to be questions on the state tests that ask students to add fractions with like and unlike denominators, and perhaps to add mixed numbers as well. Depending on how the state displays student data, teachers may be able to identify students who have difficulty with this standard, but they are unlikely to know *why* students didn't understand.

Consider, instead, what would happen if this team created a common formative assessment that would provide better information, leading to stronger interventions. Using the process of unwrapping explained in chapter 4, the team unwraps the standard and lists all of the learning targets a student needs to know to be able to add and subtract fractions, including knowing the terms *numerator* and *denominator*, understanding what fractions with like and unlike denominators are, knowing how to find the least common denominator, and understanding and being able to apply the algorithm to add fractions (convert all fractions to their equivalent fractions with a common denominator, add the numerators, and keep the same denominator). The team decides to give a short common formative assessment after teachers have taught students to find the least common denominator. Based on the results of this assessment, teachers will easily be able to group students together who didn't understand that specific learning target and provide them with more time and support. Had they waited until the end of the unit and given a test on adding and subtracting fractions, they would know which students didn't understand how to add and subtract fractions, but not why they didn't understand.

To determine what to assess, look at the learning targets your team identified for one unit of instruction through the unwrapping process. Consider which of these learning targets have the most impact on student learning. The following questions will help you decide:

- Which targets are most likely to cause certain students difficulty?
- Which targets are prerequisite skills for information to come later in this unit?
- Which targets are absolutely necessary for students to know?

Once you have identified the most important learning targets, your team must discuss the critical issue of cognitive demand—what level of thinking do we expect from our students for each learning target? Without this important conversation, your assessments might not meet the criteria for being valid.

The next step is for your team to design an assessment plan for that unit of instruction: decide how many formative assessments to give, when to give them, and which specific learning targets to include in each assessment.

Step Two: Decide How to Assess

When designing an assessment plan for a unit, your team should consider a variety of assessment strategies. Although assessment experts use different names for these strategies, they generally fit into three categories: selected response, constructed and extended written response, and performance assessments.

Selected response generally defines assessment items that ask students to select the correct answer from information provided to them. Examples include multiple-choice, matching, and true/false questions (Ainsworth, 2006; Popham, 2003; Stiggins et al., 2004).

Constructed-response (also called *extended written-response* or *supply-response*) items are those that ask students to provide their own answer to a question or prompt (Ainsworth, 2006; Popham, 2003). These include short- and long-essay responses. However, not all constructed-response questions have to include writing. For example, teachers who ask their students to complete a graphic organizer (such as a Venn diagram) are using a constructed-response assessment.

Most assessment experts also recognize performance assessments as a type of assessment strategy (Ainsworth, 2006; Stiggins et al., 2004). Marzano (2010) calls these “oral reports and demonstrations.” *Performance assessments* ask students to demonstrate their understanding of a learning target by performing in front of the teacher, who evaluates them against a rubric.

Finally, some experts discuss a fourth type of assessment in which the teacher holds a student conference to ask questions and uncover what students know about a topic. Stiggins et al. (2004) call this “personal communication,” and Marzano (2010) calls it “probing discussions.” While personal communication and probing discussions may make excellent formative assessments, they generally don’t work for teams of teachers who are gathering common data about their students because they don’t use scripted questions and won’t produce team data that can be used for subsequent planning.

Sometimes teachers are concerned that they are not using the right type of assessments—that using selected-response items, for example, is inferior to using other assessment measures. Stiggins et al. (2004) address this issue by saying that “none of these methods is inherently superior to any other, and all are viable if used well.” However, these experts do note that some assessment strategies work better in some situations than in others.

When deciding which type of assessment strategies to use, your team must consider how well each type will measure student learning of a particular learning target, as well as factors such as how quickly you will be able to get the assessment results back. In addition, because your team will develop and use these assessments, you must consider how effectively you will be able to collaboratively score assessments that rely on rubrics that might be used unevenly by different members of your team. (Collaborative scoring is discussed more thoroughly in chapter 7.)

Selected-response assessments (multiple-choice, fill-in-the-blank, matching, and true/false items) work well for formative assessments because they are easy to grade, so teams get their results back quickly. They also allow teams to use multiple measures for each learning target—in other words, several questions for each target. Even if a student misunderstands one of the questions, the assessment will still determine whether or not the student learned the information. However, experts note that it is difficult or impossible to assess very high-level thinking using this type of question (Ainsworth, 2006; Popham, 2003; Stiggins et al., 2004).

Selected-response assessments are a good choice if a team wants to quickly assess a concept and be able to respond to the needs of the students the next day. Unless the concept being assessed is complicated or involves high-level thinking, the team can design, administer, and score the assessment in a day or two.

Constructed or extended written-response items require more time for teachers to score, and the results of scoring can be uneven if teachers apply the rubric differently. In addition, students’ skill at writing can skew the results: students who have difficulty writing may appear not to have learned some targets that they have, in fact, learned. However, Stiggins et al. (2004) remind teachers that “extended

written response works well for assessing chunks of knowledge that interrelate, rather than individual pieces of knowledge assessed separately” (p. 170). If teams want to understand student thinking around a target, constructed responses are the best way to get at that information. Additionally, these types of assessments help teachers see misconceptions students have about important concepts. Math teachers often want to see the work involved when students solve a problem so that they can identify exactly what students misunderstood. Constructed-response items provide this type of information.

Many teams find that the Common Core State Standards expect students to be able to analyze and evaluate information—two high levels of thinking. Thus, they are deciding to use constructed-response items for their common formative assessments to assure they are assessing the students at the thinking level they are teaching in their classes.

Teams can use student work samples as constructed-response assessments if they are designed around specific learning targets. For example, a second-grade team gives its students blank clock faces and asks them to draw in the hands for specified times. A middle school science team provides a graphic of plant and animal cells and asks students to identify specific cell parts, and then asks them to explain how they are alike and how they are different. A high school geography team provides a map of a town and asks students to describe the relative location between two specified addresses. Any of these activities could be considered practice by individual teachers, or they could be used by teams of teachers to assess what students have (or have not) already learned.

It is important to note that when a team uses student work samples for assessment purposes, those samples must be done individually by students. A group activity or group product does not provide sufficient information about individual students and thus cannot be used as an assessment.

Students often find performance assessments to be especially engaging. These assessments ask students to carry out a process while the assessor determines, usually using a rubric, the quality of the performance. Additionally, students may be asked to create a product that will also be evaluated against a rubric. This is often the most accurate way to assess students’ ability to perform skills (Stiggins et al., 2004). However, performance assessments also take time to administer, and teams must plan ways to keep the rest of the class engaged while the teacher is watching and scoring each performance. Also, teams must develop the rubric together and discuss the nuances of its application to assure consistent evaluation.

Step Three: Develop the Assessment Plan

When developing an assessment plan, your team should consider each identified learning target, decide how to assess that particular target, and decide how long the whole assessment will take to administer.

Designing the Assessment

When developing an assessment plan, teams should consider two important factors in assessment design: making sure that the important learning targets you have identified are included in the assessment and making sure the items you write are assessing student learning at the cognitive level you identified in your unwrapping template. A quality assessment is one that includes items written around targets that the team feels are the most important and assesses at the level of thinking the team has agreed is important. Be aware, though, that in some cases, there might be more than one level of

thinking to be included around a learning target. For example, a sixth-grade science team teaching a unit on the scientific method might want to make sure that students know what a hypothesis is, that they can make a hypothesis about a specific question or problem, and that they know the steps in the scientific method. Teachers will first assess student understanding of the term *hypothesis*. On the same assessment, they will assess whether students can apply this understanding to a specific situation. The team then decides what kinds of items to use to assess each of those areas. The assessment plan in this case will look like the example in table 5.1.

Table 5.1: Sample Assessment Plan for a Sixth-Grade Team

Learning Target	Knowledge	Application	Analysis	Evaluation
Understand hypothesis, and apply it to a given situation.	Four multiple choice	Two constructed response		
Know the steps of the scientific method.	Five matching questions			

In another example, table 5.2, a third-grade team wants to know whether its students can read an informational text passage about Asian and African elephants and see that the author has used the text structure “compare and contrast.”

Table 5.2: Sample Assessment Plan for a Third-Grade Team

Learning Target	Knowledge	Application	Analysis	Evaluation
Identify and use the text structure of an informational text piece to aid comprehension.	Four multiple choice	One constructed response (Venn diagram)		

For the assessment, the teachers might ask their students to read a short essay about how these elephants are alike and how they are different. Then they might ask several multiple-choice questions about elephants that rely on the students being able to comprehend the passage. In addition, they might give their students a Venn diagram and ask students to complete it using the information in the text.

This short formative assessment allows teachers to see whether or not students can demonstrate that they comprehend the text by answering the multiple-choice questions correctly, but it also allows them to see that the students can pick out details from the text and put them in the compare or contrast portion of a Venn diagram.

While many teams will be comfortable developing their common formative assessments around only the learning targets being taught in their unit, your team may want to create assessment items that actually assess your students at a level beyond proficient so that you can measure your most able students and know which ones will benefit from additional enrichment activities during a particular unit. If this is the case, your team should consider designing the assessment with some items intended to assess a higher level of thinking beyond what you’ve identified as proficiency.

Robert Marzano (2010) describes one way to construct formative assessments so they include measuring learning targets at a level beyond proficiency. He suggests that teachers create an assessment rubric with a five-level scale: zero to four points. He describes the top level (four points) as “students make

adaptations and inferences that go beyond what is explicitly taught at the score 3.0 level” (p. 51). To assess this level, teams must add items to their assessment that are designed to assess thinking at a level higher than they would expect all students to be able to learn.

Consider, for example, the fourth-grade team at Elm Place Elementary, where teachers have been teaching to the Common Core State Standard “Explain the meaning of simple metaphors and similes (e.g., ‘pretty as a picture’) in context” (Common Core State Standards Initiative, 2010b, p. 29). In addition to the items that ask students to identify similes and metaphors from a piece of text, and to explain what those similes and metaphors mean, the teachers also added a constructed-response question asking students to develop a narrative paragraph that includes the use of both a simile and a metaphor. This item will help teachers identify students who have exceeded the evidence of proficiency and who could benefit from some additional enrichment on figurative language. In this example, the Elm Place teachers will use the additional constructed-response question to sort out their students who can benefit from an extension activity while other students continue to work toward proficiency on the learning target. Their assessment plan would look like the example in table 5.3.

Table 5.3: Sample Assessment Plan for a Fourth-Grade Team

Learning Target	Knowledge	Application	Analysis	Evaluation
Identify similes and metaphors from text.	Five matching			
Explain the meaning of common similes and metaphors.		Four multiple choice		
Develop a narrative paragraph with both a simile and a metaphor.			One constructed response	

This team would then be able to use the results of the constructed-response item to identify which students are able to go beyond the anticipated learning targets identified by the team.

Considering the Sample Size: How Many Items Do I Need?

Stiggins et al. (2004) suggest that one important factor teachers must consider when designing assessments is the number of items to include. They must be sure to include enough so that they get accurate information about student mastery of learning targets. The more items used on an assessment, the more likely the results are to be reliable—that is, to provide accurate information about a student’s knowledge around a learning target (Gareis & Grant, 2008). Gareis and Grant recommend three as the minimum number of questions or items for a learning target; they suggest that this gives a *triangulated* result; that is, the learning has been checked three times. This means that your results will be more accurate than if you relied on fewer questions.

Teams need to use a sufficient number of items to gather reliable information. They also must use formative assessments frequently. To accomplish both of these outcomes, it is important to keep assessments short. We recommend four selected-response items for each learning target if they are the only item type used. This way, a student could misunderstand or misread one of the questions and still achieve 75 percent accuracy if he or she knows the learning target. It is better practice, particularly for learning targets beyond basic-knowledge thinking, that teams use a combination of selected-response and constructed-response items.

Step Four: Determine the Timeline

Research has shown that the more frequently students are assessed, the more student achievement will increase (Bangert-Drowns, Kulik, & Kulik, 1991). However, your team must consider a variety of factors when deciding how often to administer common formative assessments. The first factor to consider is your timeline for responding with intervention/corrective instruction once you get back assessment results. For example, elementary teams with a common reading block during which students meet daily can regroup their students as often as once a week to respond to their needs as identified by assessments. In this situation, administering weekly assessments makes sense. However, if the team only has intervention/corrective instruction time for students once a week, it may make more sense for the team to administer common formative assessments every two to three weeks so that there is enough time to respond to students who need more time and support. In addition, the amount of common planning time a team has to write assessments, administer them, and analyze the data will also have an impact on assessment frequency. Keep in mind that teams that regularly do this work become more efficient than those just starting to learn the process.

We encourage teams to write and administer formative assessments at least every three weeks and as frequently as every week, with each formative assessment written around three learning targets or less. When teams use assessments with this frequency, the precision of the information they gather is much better, and therefore, their response with students is much better.

Step Five: Write the Assessment

There is a lot of information available about how to write quality assessment items. We offer some general guidelines to help you be certain your team is gathering reliable information about what students have learned.

Writing Selected-Response Items

A critical issue with selected-response items is that sometimes students who know the information answer incorrectly because they did not read or understand the question itself. When writing selected-response questions, teams should consider the following strategies to avoid confusing students:

- Include the entire question or statement in the stem of multiple-choice questions so students read through the entire statement before they begin to try to answer (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004). For example:

Which of the following strategies will work best to assess students' ability to evaluate information from text?

- a. Selected-response items
 - b. Constructed-response items
 - c. Performance items
- Use parallel construction for answer choices in multiple-choice questions (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004).
 - Be cautious when choosing vocabulary or complex sentence structure (Stiggins et al., 2004). The following is an example of an item with sentence structure that confuses the test-taker:

Having difficult vocabulary in a question will often confuse students, which will reduce its reliability, so it's important to:

- a. Include a sufficient number of questions.
 - b. Don't use vocabulary the students won't understand.
 - c. Eliminate all questions that assess vocabulary.
- Use boldface type and italics for words that students might easily miss while reading the answer stem, such as *most likely* or *best choice* (Gareis & Grant, 2008; Stiggins et al., 2004).
 - Write statements in the positive so the reader knows what the question is asking. For example, write, "Which one of the following is an example of . . ." rather than "Which one of the following is NOT an example of . . ." (Gareis & Grant, 2008; Popham, 2003).
 - Keep the list short when writing matching questions. If you need to use more than one matching set, do so (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004). For example, you could break fifteen questions into two groups—one group with seven items and one group with eight items.
 - Aim for the lowest possible reading level (Stiggins et al., 2004).

Because the purpose of the assessment is to determine what students still need help with, your team also must make sure students are unable to guess the correct answer in selected-response questions. Consider the following advice when writing these types of questions:

- Don't include "throw away" answer choices or choices intended to amuse students in multiple-choice questions. When students can easily eliminate an answer choice, their likelihood of being able to guess an answer is greater, thus making the data gathered less reliable (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004). While "eliminate the answers you're sure are wrong, and guess between the rest of the answers" is a skill students need to know for high-stakes testing, it is not something we want them to do on formative assessments.
- Make sure all answers for multiple-choice questions could possibly be correct, but that they are not so close to correct that a student who understands the content has a difficult time deciding the *best* answer (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004).
- Avoid equal-sized lists in matching questions so that students can't match all of the items they know and then guess between the leftover items (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004).
- Put answer choices in logical order (alphabetical, small to large) in selected-response questions so that students can't just guess where the answer might be. For example, you want to avoid students thinking, "There hasn't been a *b* response for a while, so I'd better choose *b* if I don't know the right answer" (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004).

When questions are multiple choice, some software programs generate an item-analysis report that teams can also use to help them plan follow-up for their assessment. The program will provide a list of which students chose each of the answers.

For example, when teams write a multiple-choice item, they might use the common misunderstandings students have about the learning target as the distracters (the wrong answer choices). By looking at the item analysis, teams know exactly why students missed that particular problem. Kopriva (2008)

explores this approach by suggesting that there are three types of distracters teams might use when writing math problems. The first is a distracter that provides evidence the student didn't understand the question being asked, the second a distracter that is the result of a common misunderstanding about the problem being solved, and the third is the result of *backsolving*—trying out each possible solution. Your team may want to consider this approach to deciding what distracters to include.

Remember that the purpose of formative assessment is to determine what students know and what they are still experiencing difficulty understanding. By keeping these strategies in mind, your team can avoid getting bad data.

Writing Constructed-Response and Extended Written-Response Items

The recommendations for making selected-response questions easy to read and understand apply to constructed-response questions as well. However, the experts have some additional recommendations.

With constructed-response and extended written-response questions, it is important to give students a context to use for their answer (Gareis & Grant, 2008; Popham, 2003; Stiggins et al., 2004). For example, if students are asked to write about the effects of environmental issues on the future of the United States, remind them about specific issues, such as the 2010 Gulf of Mexico oil spill or the greenhouse effect created by car emissions. This helps them to frame how they will answer. The resulting question might be:

Environmental issues can have social, economic, and political effects on our country. For example, the 2010 oil spill in the Gulf of Mexico affected our economy through the tourism industry—an economic effect. The greenhouse effect has resulted in many new laws reducing car emissions—a political effect. Choose one environmental issue that has had such an impact and explain what the social, political, and economic impact was.

In addition, make sure the question or prompt is novel—that is, make sure it isn't about a topic that you've already discussed in class (Gareis & Grant, 2008; Popham, 2003). For example, if you ask students to describe the effects of the Civil War on the economy of the South, and this was a topic of discussion during the unit, they are only being asked to remember what was discussed. However, if you ask them to describe how the economy of the North was affected (and this hadn't been discussed in class), the assessment determines whether students are able to understand the reasoning around this concept by applying it in the new circumstance.

Finally, make sure that the directions for the assessment are clear and easy to understand, and leave space for your students to fully answer the questions.

Designing Performance Assessments

Stiggins et al. (2004) recommend teams use performance assessments when the learning target is a skill that can best be assessed by observing the student perform that skill. He notes that with performance assessment, students should know how they will be evaluated (the scoring rubric) during their performance.

Students should not be given a choice between different performance tasks unless all tasks are the same level of difficulty and designed to assess the same learning targets. Otherwise, students could choose to do the task they know and understand and skip a task they don't know (Stiggins et al., 2004).

Additional Guidelines for Writing Assessment Questions

When writing assessments, teachers will often look for pieces of grade-level text they can use to assess reading comprehension. Remember that common formative assessments are generally written around specific reading skills or strategies. Therefore, the text must provide the appropriate opportunity to assess that strategy. For example, if the team has taught the text structure *cause and effect* and wants to see if students can analyze a piece of writing and use it to aid comprehension, then it must find a grade-level appropriate piece of text written in the cause-and-effect format. Textbook assessment materials and released materials from state tests (found on state websites) are places to look for text passages to include in assessments. In addition, appendix B in the English Language Arts Common Core State Standards provides a variety of suggested text pieces to use for assessing students on these standards.

Sometimes teams want to use questions from test banks, their textbooks, or other sources. When using these questions, teams must take some important steps to make sure these sources will yield accurate information about student learning. It is important that all questions are aligned to the learning target being assessed, and questions must also be written at the level of rigor set for that target. For example, if the learning target being assessed is that students can draw conclusions from data gathered in an experiment, it is important that the question isn't a simple knowledge question. Instead, students should be given a data set from an experiment and asked to write and explain what their conclusion is. In addition, it is very important for teams to consider the quality of the items they are selecting—just because they are provided by an outside source does not guarantee they will measure the learning targets identified by the team or that they are well written.

During the process of writing assessment questions, teams might wonder if they should use certain types of questions or structure them in a certain way to prepare students for high-stakes tests. Teams should discuss this issue to determine how to include these important strategies into their curriculum if they decide it's appropriate. Instead of using formative assessments for this purpose, many schools and districts use the benchmark assessments that measure student progress over a period of time to provide students with practice in test-taking strategies. These assessments are often written using items similar to those on the state test. Once the assessment has been given, the teachers can go back over the items to show students test-taking strategies they can use in the future.

One of the benefits of working collaboratively is that teams become better item writers and assessors by sharing ideas and information. Teams often want to know how to get started on this process and how to be as efficient as possible. Our first recommendation is that, at least early in your work, the entire team works together to plan and write the assessment. The value of learning together often outweighs the time it takes to create collaboratively. In fact, as teams get started they may take more time to design and write their assessments and find they can't complete as many assessments as they'd like. We believe that this initial learning process makes quality more important than quantity!

We do suggest, however, that teachers bring any assessments they've used in the past for consideration—you don't have to reinvent the wheel if it's already working. These assessments provide a starting point for the team's work. However, it is really important that the team completes the assessment plan first before choosing or writing items. Then, as the team writes the assessment, each teacher must make sure that students will understand and be able to read the included items. These multiple points of view will make it easier to spot confusing items.

We also strongly recommend that teams keep feedback from their already-administered assessments about what worked and what didn't. The following year, the team can then update the assessments rather than starting from scratch.

Step Six: Review the Assessment Before Administration

It is important for your team to review the assessment closely prior to giving it to students. Are the directions clear? Do students know what you are asking them to do and why? Some teams actually list the learning targets on the assessment itself, followed by the questions related to each target. This helps students to clearly understand the context of the questions. Teams should also review the expectations about how much time it will take to administer the assessment.

Step Seven: Set Proficiency Criteria and Decide How to Gather the Data

During this step, teams first discuss what proficiency will look like. For selected-response questions, decide how many correct responses students must have for each learning target. For example, if there are four multiple-choice questions, students who get at least three correct will be considered proficient. A scoring rubric is necessary to determine proficiency for constructed- and extended written-response items. For example, your team might decide that on a six-point rubric, a level four will be proficient.

There are two types of rubrics: holistic and analytic. A *holistic rubric* is used to score the overall proficiency of a student work sample or performance. An *analytic rubric* is used to evaluate each criteria or trait of the student work sample or performance separately to analyze student learning in a more specific way (Arter & Chappuis, 2006). The advantage of using an analytic rubric for a common formative assessment is that you will be able to determine exactly what areas of learning individual students need additional time and support in—as long as the rubric was designed around the specific learning targets. Chapter 7 discusses creating rubrics in greater detail.

In the next chapter, we'll discuss how teams can use their data to design units and pacing guides. Before teams can use their data for the bigger picture, though, they must gather and organize it. Because data analysis is done by learning target first, that is the structure by which the data must be gathered. Then your team will look at how each student did on each learning target. Teams can save time during their data team meeting by having a plan for how to gather data, as shown in table 5.4.

Table 5.4: Plan for How to Gather Data

Target	Students Who Need More Time and Support	Students Who Will Benefit From More Practice	Students Who Will Benefit From Enrichment or Extension
Target One			
Target Two			

Teams need to make sure their assessments are both valid (assess the expected learning targets at the anticipated level of thinking) and reliable (provide accurate information about what the students have learned). The Evaluating the Quality of an Assessment tool on page 106 of the Tools for Teams appendix will help you assess the quality of your assessments.

The Sample Protocol for Developing an Assessment tool in the Tools for Teams appendix on page 107 outlines the seven-step process for developing an assessment, highlighting the main tasks of each step. Your teams can use this tool for reference as you begin the process of developing your own common formative assessments. The team uses the Assessment Plan document on page 108 as it begins the process to decide which targets are being assessed and what the expected level of thinking is for those targets. It then completes the table by deciding what types of assessments to use and how many items it will need to include.

Final Thoughts

By thoughtfully designing assessments around specific learning targets and gathering and analyzing the data, your team will know exactly what your students have learned and what they are still experiencing difficulty understanding. This allows you to develop a strong instructional response to assure that the additional time and support you give to students will meet their learning needs. In chapter 6, we discuss creating pacing guides and designing units using the information you've gathered through the assessment process.



CHAPTER 7

Now What? Using Data to Make a Difference

KEY POINTS

- The most important result of using common formative assessments is the response teams develop and implement to support student learning.
- In a common formative assessment, data must be gathered *by learning target* for each student.
- Data conversations can follow a prescriptive protocol that allows teams to make effective use of their time and create a safe place to share ideas and information.

While they offer many benefits, the primary goal of common formative assessments is to provide information about student learning and to identify which students are in need of additional time and support. In fact, they are an integral part of a school's Tier 1 instruction within the response to intervention model (Buffum, Mattos, & Weber, 2009). Specifically, in schools that operate as PLCs, teams use the results of their common formative assessments to identify students who have not reached proficiency on prioritized learning targets. What differentiates teams in PLCs from traditional teams is the response itself. In PLCs, when students are identified as not yet reaching proficiency on the skills and concepts considered essential, there is a collective and systematic first response *within the team* designed to provide immediate support. This differs greatly from traditional teams in which members feel it isn't their responsibility to provide support. These teams respond to student needs simply by looking for other teachers or staff, such as specialists, to provide the intervention students need.

In addition, many teams get stuck when examining data and planning a response. They get stuck because they are trying to use a process to analyze their data that is cumbersome and time consuming. For example, many data-analysis protocols ask teachers to first look for facts, then create a hypothesis, and so on. Teachers are taught to look for patterns in their data by graphing their results over time. While these processes are effective when examining summative data, they are less likely to help teams use common formative assessment data effectively.

Mike Schmoker (2003) reminds us that we don't need "sophisticated data analysis or special expertise" (p. 23) to collect and use the data we need. The step-by-step process we describe in this chapter will help your team make decisions about student learning and plan intervention based on meaningful and timely data—your common formative assessments. This process allows teams to plan these next steps in a way that will be practical and will make a difference for students. The Tools for Teams appendix contains two tools to help teams plan as they progress through this process: the Data Team Meeting Template (page 112) provides several graphic organizers for teams to use when bringing their data to the table, and the Protocol for Data Team Meeting document (page 115) highlights the steps in the process.

Step One: Gathering the Data

In chapter 5, your team was encouraged to design formative assessments around a small number of learning targets. You created an assessment plan listing the specific types of questions you would ask and how many you would ask for each of the targets being assessed. So when gathering the data from the assessment, your team should first list the item numbers from the assessment that were written for each learning target. For example, questions one to four were written to assess learning target one, questions five to eight to assess learning target two, and questions nine and ten to assess learning target three.

For example, the seventh-grade math team at Washington Middle School developed a common formative assessment to determine which students understood adding integers. Its assessment plan is shown in figure 7.1.

Learning Targets	Knowledge	Application	Analysis	Evaluation	Total Items/Total Time
T.1: Add negative integers		Four multiple-choice questions (one through four)			Four (six minutes)
T.2: Subtract negative integers		Four multiple-choice questions (five through eight)			Four (six minutes)
T.3: Apply to problem		Two constructed-response questions (nine and ten)			Two (eight minutes)

Figure 7.1: Seventh-grade assessment plan for integers.

To make things simple, the team wrote its assessment so that questions one through four assessed target one, questions five through eight assessed target two, and questions nine and ten assessed target three. Members decided that students would have to get three out of four questions correct to be proficient on targets one and two, and students would have to score a three on a four-point rubric for each of the questions in target three.

At the data team meeting, team members brought with them a list—by learning target—of which students have not reached proficiency, which have reached proficiency but would benefit from additional

practice, and which students are beyond proficiency and would benefit from enrichment. This allowed the team to determine the total number of students in need for each level of intervention or enrichment.

In our example of the math team, each of the teachers brought his or her collected data in a form similar to the one in figure 7.2.

Learning Target One: Solve a Problem		
Needs Time and Support	Perhaps More Practice	Ready for Enrichment
Sally	Matthew	Aaron
Mary	Jeremy	Andrew
Josh	Samantha	Max
Jennifer	Erin	Cindy
Learning Target Two: Solve a Problem		
Needs Time and Support	Perhaps More Practice	Ready for Enrichment
Sally	Jennifer	Aaron
Jeremy	Mary	Andrew
Josh	Samantha	Max
Matthew	Erin	Cindy
Learning Target Three: Apply a Problem		
Needs Time and Support	Perhaps More Practice	Ready for Enrichment
Sally	Josh	Max
Mary	Jeremy	Cindy
Matthew	Aaron	Erin
Jennifer	Andrew	Samantha

Figure 7.2: Sample of collected assessment data for three learning targets.

Step Two: Analyzing the Data

Teams often want to begin planning their next steps for instruction right away when they first meet to look at the data. However, your team will be able to develop a better response plan by taking several initial steps. First, review the proficiency level you established for each learning target to determine if it was set at the appropriate level. For example, in the Washington Middle School example, the team reviewed whether or not level three on the constructed-response rubric represented an appropriate proficiency level for that learning target. Team members also looked to see if any of the multiple-choice questions appeared to be of concern. Sometimes most students who appear to understand the information (based on how they answered other questions for a particular learning target) answer the same question incorrectly. Other times students who appear not to understand a learning target get a particular question correct. Teachers may infer from this observation that a question was not well written. Your team should make note of any problematic questions under the heading “Which questions need to be reviewed?” on page 114 of the Data Team Meeting Template in the Tools for Teams appendix.

Once you note any assessment items that need additional review, your team can begin to analyze the data. Start by listing each learning target and the total number of students who scored below proficiency, met proficiency, and exceeded proficiency. During this step, focus on determining the total number of students who will need additional time and support—not on how that support will happen.

Sometimes schools and districts purchase computer software that score selected-response questions and generate reports with the data by student and learning target. If not, teams can create their own spreadsheets that allow them to sort the data from a formative assessment by learning target (table 7.1).

Table 7.1: Data Sorted by Learning Target

	Learning Target One Proficiency: $\frac{1}{4}$ or 75%	Learning Target Two Proficiency: $\frac{1}{4}$ or 75%	Learning Target Three Proficiency: $\frac{1}{4}$ or 75%
John	75%	75%	3
Jeremy	100%	75%	4
Caitlin	50%	100%	2
Missy	75%	100%	4
Jason	100%	100%	3
Brian	75%	25%	2
Sarah	50%	75%	4

As you discuss your results, you can re-sort for each learning target to see which students scored at each level.

After looking at the total number of students at each level, consider some more general results. Did the same students seem to miss all of the learning targets? Is the number of students who need additional help only a small percentage of the total number of students, or is there a larger majority who didn't understand a target? The answers to these questions will help your team decide whether or not additional whole-class instruction is needed around a learning target, or if teachers can respond to students in smaller groups.

Step Three: Planning the Response

Once your team knows the total number of students in need of additional time and support for each learning target, you can begin to plan how to respond. To do this, you must further analyze the results of the assessment. The following questions will help your team effectively uncover a plan to respond to the data.

Which Teacher Was Most Effective?

Your team might recognize that one teacher was more successful on a particular learning target than the rest of the team. One of the benefits of common assessments is that teams can compare their results to the results of their colleagues. In fact, according to DuFour et al. (2010a), each teacher needs to see the data for his or her students “*in comparison to other students in the school attempting to meet that same standard*” (p. 189). They remind us that “without relevant information on their respective strengths and weaknesses, teacher conversations regarding the most effective ways to help students learn a concept will

deteriorate into sharing of uninformed opinions—“This is how I like to teach it” (DuFour et al., 2008, p. 27).

A teacher who was more successful at teaching a particular learning target can share his or her instructional strategies with the group, and the team might decide that students who need more time and support should be assigned to that teacher to receive further support on that target.

What Mistakes Did Students Make?

Your team might also discover when analyzing data that each teacher has roughly the same number of students who experienced difficulty on a learning target. In this case, using one teacher’s instructional strategy for additional support is not necessarily the answer. Instead, create a hypothesis as to why students experienced difficulty in understanding the target. This will help plan how you will support students with additional instruction.

Start by listing possible reasons why you think students experienced difficulty. Teachers might, for example, notice that the students who experienced difficulty appear to have weak background knowledge or prerequisite skills. They know that these students will need additional scaffolding and reinforcement of background knowledge to learn the target. Teachers might also hypothesize that students had difficulty because the target is an abstract concept, thus students will benefit from more concrete learning opportunities around this target. Teachers might discover that certain students have a common misconception that is getting in the way of their understanding of a learning target, so students will need help to overcome their misconception.

For example, the physics team at Harbour View High School gave its students an assessment asking them to analyze some graphs from an experiment and then develop and explain a conclusion drawn from the information. The team members used eight multiple-choice questions and two constructed-response questions on their assessment. As they looked at their results, they discovered that a number of students got all of the multiple-choice questions correct, but missed the constructed-response questions asking them to develop a conclusion. The team hypothesized that these students really did understand how to analyze information from a graph, but they didn’t know how to develop a coherent conclusion *in writing*. By creating hypotheses about why students didn’t learn a particular target, a team builds possible ways to respond with corrective instruction.

In chapter 5, we discussed that some teams write their multiple-choice questions with specific distracters to know what mistakes their students are making. If your team wrote questions this way, it will be important at this step to look at the item analysis to see which students made which mistakes. For example, if the multiple-choice question has four choices (one correct and three based on common misunderstandings students have), you can know more specifically why students got it wrong.

Since the purpose of formative assessment is to know exactly which students haven’t learned a particular target and *what we need to do next for them*, it’s important to use the data from our assessments in an unbiased way. However, while teachers need to be open-minded about why students are experiencing difficulty learning a particular target, it is important that they use their knowledge about students to create a meaningful hypothesis about why students didn’t learn the target. This prevents us from using the “they just didn’t even try” excuse for students who regularly experience difficulty meeting proficiency.

How Can We Provide Corrective Instruction?

Once teachers brainstorm their hypotheses, they then turn to plans for providing corrective instruction based on why they think students did not reach proficiency. In the example from the Washington Middle School math team, the team discovered that subtracting negative integers is a concept with which many of its students had difficulty. Knowing that this is a highly abstract concept, members hypothesized that students really didn't understand the abstraction. Therefore, they planned to reteach the concept using manipulatives to make it more concrete.

One teacher shared a strategy using chips with different colors on each side. The red side represents a negative number and the yellow side a positive number. In the math problem $3 - (-2) = x$, the student lays out three yellow chips to represent the first number in the problem. He then adds two yellow chips and two red chips because together they equal zero. (The student knows that you can add zero without affecting the outcome.) Then the student removes two red chips (representing that he has subtracted a negative two). The final solution is five because the student has five chips remaining. The teacher then reviews the rules for subtracting negative numbers to demonstrate that the subtraction sign and the negative sign cancel each other out so that the final sum is the addition of the two numbers.

In the earlier example of the high school physics team, the team responded to the students who got the multiple-choice questions correct and the constructed-response questions incorrect by helping students learn how to develop a strong paragraph in writing explaining their conclusion.

What Other Resources Do We Have for Support?

Sometimes after examining data, a team realizes that no teachers used a highly effective teaching strategy for a particular learning target, and the team can't explain why its students didn't do very well. In this case, the team members decide that their next step should be to go back into best-practice research to see what additional strategies they might want to consider. For example, Marzano, Pickering, and Pollock (2001) describe nine highly effective, research-based strategies to help students:

1. Identifying similarities and differences
2. Summarizing and note taking
3. Reinforcing effort and providing recognition
4. Homework and practice
5. Nonlinguistic representations
6. Cooperative learning
7. Setting objectives and providing feedback
8. Generating and testing hypotheses
9. Cues, questions, and advanced organizers

Teams can choose one strategy to use in their intervention efforts with students who didn't reach proficiency. Reviewing the professional literature to determine the most current thinking about best

practice will help the team build a repertoire of strategies that it can use to teach a concept or learning target. The original strategy team members used in their classrooms might still work for the majority of students, but it hasn't worked for all students. This step asks teams to consider additional strategies beyond their team.

Step Four: Reviewing the Assessment

Your team should also review the assessment as a whole so you can discuss whether or not you want to use the same assessment items in the future. Look back at the questions you identified as in need of review. Some questions might just need to be reworded because they are not clear enough, and others, such as constructed-response items, might need a clearer context. You might decide that some of your items weren't written at a high enough level to assess the thinking your instruction was designed to target. Your team should also discuss how long it took for students to take the assessment, and how long it took you to score their responses. Finally, your team should discuss instructional and pacing issues it might want to adjust in the future.

Consider, for example, the sixth-grade English language arts team at Meadowview Middle School. Members realized during their data meeting about an assessment on sentence structure that the questions they were asking were primarily factual questions and that students already knew the material they were assessing. The teachers realized that instead of asking questions about types of sentences, they would get better information if they asked students to take a piece of writing and edit it to show how to use a variety of sentence structures (combining sentences and varying sentence length).

Or consider the biology team members at Jefferson High School who realized after their assessment that the majority of their students were still having difficulty understanding the difference between *diffusion* and *osmosis*. They agreed to spend two more days of instruction and use several additional activities. They noted this on their pacing guide so they would be sure to do the same the following year.

At Earhart Elementary School, the third-grade team members questioned the time it was taking to administer their reading assessments. After much discussion, they decided that they were trying to assess too many learning targets every time they did a formative assessment as they were trying to assess some vocabulary and some strategies, as well as including some writing targets. Instead, they agreed that they would assess one or two skills only—such as finding the main idea and details—and that they would base their questions on one reading passage.

Your team should also review the consistency of your scoring practices in this step. When a team uses a rubric to score student work products or performances, it is important that all teachers apply the same criteria to those products and performances. To test whether or not teachers are applying the same criteria, teams can do the following exercise: Each teacher brings three or four random samples of student work to the group. Each teacher then scores each sample and puts his or her score in a private location (on a scoring sheet that isn't passed along with the sample, on a sticky note placed on the back of the sample, and so on). Once each teacher has scored all the work samples, the team looks at the range of scores given to that piece of work. If the scores are all within a point or two (depending on the number of points on the rubric), then the team can assume individual members are applying the rubric equally. However, if the sample has a range greater than one or two points, it is important for the team to discuss the disparate score to come to consensus on the appropriate mark. This process not only validates the scoring for the work samples brought to the meeting but for the rest of the student work as well.

Another process some teams use to ensure consistency is to score all of the work together at their meeting. In this case, teachers exchange work samples so that they're not scoring the work done by their own students. Before they begin the scoring process, the teachers examine several anchor papers that represent the quality of work for each level of the rubric.

Step Five: Next Steps

Once you have analyzed your data, you will develop a plan for your team's next steps. You have shared ideas about the best instructional strategies for intervention and how you will respond to students who have not learned. Now you need to decide who will deliver those corrective instructional strategies and to which of your students. This begins with finding the time.

Finding Time to Respond

Often the decision about how to respond is dependent on the school's master schedule or a team's schedule. In this next section, we'll look at some specific ways schools and teams have reworked their schedules to find time to respond.

Elementary Teams

Many elementary schools create their master schedules to provide common teaching time for all teachers at a particular grade level. This time might be called *walk-to time*, *intervention time*, or even just *the literacy or math block*. What's important about this common teaching time is that all teachers on the team share their students, so that teachers can move students into different classrooms depending on students' specific needs. During this time, all of the teachers are working with groups of students to respond to assessment data, whether to support learning needs, practice, or enrich. None of the teachers is providing new instruction during this time, so none of the students misses any important initial instruction, which would put him or her even further behind.

For example, at Jackson Elementary, three first-grade team members just finished their instruction about how to tell time to the nearest half hour using an analog clock. They used a short formative assessment to determine that approximately 20 percent of their students will need some corrective instruction before the class can move on to the next lesson in the unit. During its data meeting, the team created a plan to spend an additional two days to make sure that all students learned this important learning target. Teachers decided to divide their students into three groups, with each teacher responsible for one group.

One teacher will work with students who need corrective instruction using a strategy the team members developed during their data meeting based on why they think the students didn't learn the target. A second teacher will take a group of students who passed the assessment but will benefit from some additional practice. The third teacher will take the students who know the target and will benefit from enrichment. The students who didn't learn the target will also be given another assessment to make sure they've learned it after receiving additional support. This assessment can be a different formative assessment that assesses the same target, or it might even be the same assessment given earlier. If there are students who still need additional help after this second assessment, the team will plan a Tier 2 intervention for this small group.

Sometimes elementary teams add additional staff to help keep their student groups smaller. For example, at Skyview Elementary, the first-grade team planned a response similar to that of Jackson

Elementary, except it has a math specialist and an instructional aide available during the scheduled math time. By adding these two staff members, the team is able to create two smaller groups of students from those who need corrective instruction: one taught by a classroom teacher and one by the math specialist. The rest of the students are divided between the two remaining teachers and the instructional aide for additional practice (taught by one of the teachers and the instructional assistant) and for extension activities (taught by the third teacher).

Secondary Teams

Middle and high school teams often have more difficulty finding time for corrective instruction because their master schedules are so complex. However, there are still a number of ways schools can find the time they need—they just need to be creative. For example, one high school we've worked with developed its master schedule so that every time one of its core ninth-grade classes meets, there is at least one other section of the same core class meeting at that same time. In other words, if Algebra I is taught during first period, there are at least two sections meeting first period. This allows those two teachers to exchange students after formative assessment to provide intervention and enrichment.

Many middle schools have two or more teams at each grade level, so their master schedules make it possible for teachers to exchange students between the teams when responding to formative assessments. One caution we make is that when teachers exchange students, they do so for a day or two only at one time based on the results of a formative assessment; they should not group students permanently or for a long period of time. This flexible grouping will support specific learning needs without exposing less-able students to the negative outcomes of grouping them into a "low group."

In addition, many middle and high schools are creating master schedules that have a built-in intervention period either daily or several days a week. For example, if the school has an eight-period day, a few minutes are shaved off each period to create a half-hour period during the day when all students are available for intervention or enrichment. Then the team can organize the students who need more time and support into smaller groups. Students who already learned the target receive enrichment during this time.

Finally, some high schools have resource rooms for students to use during their study halls. If a student is identified as not having reached proficiency on a particular learning target, he or she can go to the resource room to get help from a staff member who is not assigned to be teaching at that time.

Designing Tier 2 Response

No matter what structure your team uses to provide time for response, the most important thing is that there is a response and a follow-up assessment for any learning target a student missed. Teams must decide when to give follow-up assessments and what they will look like. As we mentioned, some teams choose to use a different version of the first formative assessment; others use the exact same assessment. The critical issue is that teams measure whether or not students have mastered the target after the intervention.

So what happens when students still haven't mastered an important learning target after the initial instruction, formative assessment, and corrective instruction? What happens when a student takes the end-of-unit test and doesn't show proficiency on one or more essential standards? Even when the initial core instruction is over, teams within PLCs don't abandon students who haven't yet mastered essential

learning outcomes. In a PLC where the school has developed a tiered system of response—a pyramid of interventions—there are many ways students can receive more time and support to learn beyond the core instruction. These are known as Tier 2 and Tier 3 responses. These additional learning opportunities are provided to the smaller number of students who need even more support, and often happen in addition to new instruction in the classroom. For example, in an elementary school where a student received initial instruction on a reading concept, was identified through a formative assessment as needing extra help, was provided with correction instruction, and then was identified in a follow-up assessment as still not having mastered the target, the student is given more intensive smaller group intervention. He or she may be pulled out for small-group instruction or a specialist might “push in” to the classroom. Either way, the student continues to get instruction and support on the essential outcomes that are most important for learning. Visit “Evidence of Effectiveness” at www.allthingsplc.info for more examples of schoolwide systems of intervention and enrichment.

Creating Safe Conditions for Teams

When teams examine the results of common formative assessments together, it is important that team members create a safe environment for one another. However, as Wellman and Lipton (2004) point out, “safety and comfort is not always the same thing in collaborative settings” (p. 12) and teams must work on developing trust around their use of data. If you never discussed your team norms for behavior concerning the use of data, you might find you need to revisit this process after your first assessment or two. Norms should include the understanding that teams will treat the information from their assessments objectively. Teams shouldn’t be judgmental—rather, members should examine data and state facts so that they can create accurate hypotheses about why students struggled (Steele & Boudett, 2008).

The Best Hopes/Worst Fears activity (Wellman & Lipton, 2004) is one way for teams to establish a trusting environment. In this activity, members brainstorm on a T chart their best hopes and worst fears about using common formative assessments. Team members often acknowledge that they are afraid their colleagues will be judgmental, that their results will not be as positive as they had hoped, that they will look bad to their teammates, and so on. Once these concerns are out on the table, your team can address each of them by creating data norms, such as the following:

- We will look for facts, not blame.
- We will learn from the results of our assessments.
- We will not judge our colleagues.

Once you’ve discussed your fears, remind yourselves about why this work is so important—what are your best hopes? These hopes will likely include higher student success, more personal satisfaction, increased knowledge of instructional strategies, and various other learning outcomes. You’ll likely conclude that the benefits of common formative assessments outweigh the possible fears.

Final Thoughts

It is critical for teachers to know what their students are learning and not learning so they can make decisions about instruction. By doing data analysis together, your team will be able to provide much stronger and more successful intervention and enrichment, as well as gain valuable professional development for all members. But remember, the best data analysis leads to action by the team—otherwise students will not benefit.