BRING

SAMPLES FROM YOUR SCHOOL – (single copy or online access to)
One to two samples (a single copy of each if you print) from your school of each of the following:

- Your school’s assessment policy (if no policy, any guidelines for teachers about assessment)
- Examples of data that your school collects on student learning (e.g. report card data, external tests results, running record data, etc.)
- A page from a grade or record book
- Report card from any section(s) of your school

ASSESSMENT LEADERSHIP PROCESS AT YOUR SCHOOL
While we will be building understanding, skill, and knowledge on assessment practices, the emphasis is on LEADING the assessment process in a school.
We will address these two Standards for International Principals:
ACCOUNTABILITY FOR LEARNING An international school principal’s leadership results in student progress consistent with the school’s mission and curricular goals.
CURRICULUM AND ASSESSMENT LEADERSHIP The international school principal leads a curricular system based on the most recent understanding about learning, including developing, implementing and monitoring curriculum and learner progress.

- If you are a principal, write a brief description of what your current role is with regard to student assessment. Include structures and processes, including teams, etc. that support assessment of learning.
- If you are not a principal, interview a principal in your school and bring some brief notes on your principal’s perception of his/her role in assessment.
- For both cases above, also briefly state what is the BIGGEST ASSESSMENT CHALLENGE at your school.

Please either print the above OR have access to a digital version.

READ

Please read the following article and make some brief notes.
- A Principal’s Guide to Assessment by Chris Jakicic (included below)
PREPARE

☐ AN ISSUE OR PROBLEM
In preparation for the small group sessions in the afternoon please bring a written (one page) case study or example, dealing with a problem or situation you have confronted during your professional practice.

☐ JOIN SCHOOLOLOGY: ...your required online portal for everything related to your PTC course, including all necessary course materials, which will be available about 10 days prior to the start date of your course. Register for Schoology using a personal email (NOT a school email) that you will always have access to.

- Have a Schoology account? Click here.
- New to Schoology? Click Here.

Use the access code below for your session and location to join your class GROUP: Do NOT click “course”, click “group.”

- Schoology Access code for Miami 17 -23 June: J9PQM-MJ4DT
- Schoology Access code for London 3- 9 July: CWJZ5-MVQFZ

☐ WATCH
...a short video to prepare for your course.

☐ BRING A LAPTOP (REQUIRED)
- Your course materials are digital. Hard copies of the materials will not be available on site. Bring a laptop computer with appropriate adapters and wireless internet capability. Click here for more information about iPads and tablets. Chromebook users can use the free Kami extension to annotate.
- Download Adobe Acrobat Reader DC now in order to be able to access your materials properly.

We look forward to our learning together!

Miami
Bambi Betts

London
Melissa Schaub and Lauren Mehrbach
A Principal’s Guide to Assessment
Chapter Three
By Chris Jakicic

All educators are facing tough scrutiny about student achievement issues in this era of high-stakes testing. Principals often wish that someone would hand us a guidebook explaining how we can truly ensure that all of our students will learn. In the absence of such a guidebook, we must look at the research about what successful schools have done to improve their achievement levels and decide how we can apply these practices to our own schools. With this in mind, it would be hard to ignore the compelling research about the benefits of using assessments to guide our instructional practices as we work to help all students learn at high levels.

One of the most cited pieces of research about the effects of assessment is summarized in Paul Black and Dylan Wiliam’s article “Inside the Black Box” (1998). The authors examined over 250 international research studies about schools successfully improving student learning and concluded that schools could expect a .4 to a .7 standard deviation increase in student achievement if teachers implement formative assessments in classrooms. This level of improvement is significant enough for any principal to want to duplicate those results.

This chapter explores what principals need to know to support teachers as they create and use an effective assessment system that includes formative assessment. I will also consider problems that might emerge as teachers use new assessment practices and suggest solutions.

What Is a Balanced Assessment System?

One of the first considerations when evaluating the effectiveness of your assessment practices is whether teachers and administrators get the information they need as a result of the assessments currently in use. Teachers want to know which of their students are learning and, if some are not, what they need to do to help those students. Principals want to know if there are curriculum, instruction, or pacing issues that must be changed to create a culture in which all students are able to learn. In order to obtain all the information, they need, educators need a balanced assessment system that includes both formative and summative assessments.

The National Center for Fair and Open Testing (2007) defines summative assessment as “the attempt to summarize student learning at some point in time, say the end of a course” (p. 1). Teacher-created final exams, achievement tests, and state tests are all examples of summative assessments. Summative assessment data are used to give grades, to determine whether the curriculum and instructional practices being used are effective, and whether the pacing provided appropriate time for students to learn the material being taught.

James Popham (2008) defines formative assessment as “a planned process in which teachers or students use assessment-based evidence to adjust what they’re currently doing” (p. 6). These assessments can include quizzes, worksheets, or other classroom activities that provide feedback to teachers about whether students have learned a particular learning target—while they are still teaching that unit. Formative assessments are written around learning targets that teachers have identified as the specific knowledge and skills their students need to understand a standard. For example, in science, students are taught how to use the scientific method to solve a problem. Teachers who are familiar with this unit know that the concept of control and variable is more abstract than other concepts in the unit, so they plan to do a short assessment after teaching the concept of control and variable to see whether the students understand it. With the results of their formative assessment, the teachers know which students need more help to attain this learning target.

Rick Stiggins, Judith Arter, Jan Chappuis, and Steve Chappuis (2004) use the term assessment of learning instead of summative assessment. Stiggins defines assessments of learning as “those assessments that happen after learning is supposed to have occurred to determine if it did” (Stiggins et al., 2004, p. 31). He explains that assessments for learning (formative assessments), by contrast, are intended to provide information to teachers and students during the learning process. No matter what terminology is used, researchers clearly acknowledge that teachers need access to both types of assessment data.

Even more recently, some researchers are writing about another type of assessment, what Kim Marshall (2008) is calling interim assessments, also known as benchmark assessments. These assessments are given every four to nine
weeks to provide teachers with information about whether their students are making progress toward mastery of their grade-level standards. These assessments are often developed by a team of teachers and are used to discuss how to respond to students who are not learning, to ensure that students are retaining the information they have learned, and to predict which students will need more support to attain mastery of grade-level standards.

**Assess Your Assessments**

The first recommendation for any school evaluating their current assessment practices is to create a list of all the assessments currently being given to students and note who develops the assessment and who uses the information. Then, the teachers and principal together should examine the list to see what areas are not covered adequately by current assessments and what areas are overcovered through assessments that provide duplicate information. If there are areas of duplication, consider eliminating one or more tests. Alternately, use one of the duplicate assessments for progress monitoring of only those students who are experiencing difficulty so that teachers know if their supports are working.

In my own school, Woodlawn Middle School in Long Grove, Illinois, we learned that we had multiple summative assessments for our students in many areas, but few formative assessments. For example, we were using the Measures of Academic Progress (MAP) twice a year, as well as quarterly benchmark assessments in reading. In addition, of course, we had the results of the state assessment. Teachers believed we were doing too much testing; however, we also realized that we didn’t have any formative assessments to tell teachers what they needed to do the next day in their classrooms. After examining our assessments, we decided to add formative assessments and to eliminate one of the times during the year that we administered the MAP test.

An important decision schools must make is whether teachers will establish “common” assessments designed and administered by teams, or whether each teacher will develop and use his or her own assessments. On one hand, developing assessments as teams creates discussion about what is important to teach and what mastery will look like; this helps to create the “guaranteed and viable curriculum” that Robert Marzano (2003) tells us is the most important factor for successful schools. Teams of teachers can then share effective instructional strategies and can respond to student needs collectively. On the other hand, teachers who work independently can accomplish the work more expeditiously because they don’t need common planning time. They’ll still need time to write good assessments, but teachers who teach the same subject or grade level won’t all need time together.

Principals must facilitate the decision about whether teacher assessments will be common or individual early in the process so that teachers know the expectations upfront. Once teachers have moved forward with writing their own assessments, it may be much harder to backtrack and create team assessments. If the school decides that the assessments should be common, it will be important for the principal to create a schedule that allows for common planning time or consider how to use staff development time for assessment writing.

**Support the Development of an Assessment System**

Larry Ainsworth and Donald Viegut (2006) lay out the support teachers will need as they work through the process of developing assessments: professional development, time for writing assessments and analyzing results, clerical help, and opportunities for coaching and discussion around assessment. While all of these needs apply to either individual or common assessments, their application will be different depending on which process a school uses.

For example, initial staff development about how to write and use good assessments will be the same in either case, but follow-up coaching and support will be very different. If teachers use common assessments, they will be able to help each other analyze and respond to the data. If teachers are working individually, the principal will want to assure that all teachers are being coached about how to effectively use the data. When thinking about coaching individual teachers, consider whether the teacher has enough experience and background to effectively analyze the data. For example, Roberta Buhle and Camille Blachowicz (2008/2009) describe a coaching session with a kindergarten teacher who saw that a pre-assessment she had given indicated that most of her students knew the names of the letters, and yet still intended to follow the typical plan to teach a letter-of-the-week. The coach intervened to point out the flaw in this thinking. This example shows how important it is for the principal or the coach to be on the lookout for teachers who need more help understanding how to respond to their data.
**How Frequent Is Frequent Enough?**

Once teachers begin working on their assessment system, the next question that will likely emerge is, How frequently must we assess to collect the information we need to make a difference? At Woodlawn, we started by developing what we now call benchmark assessments, which were given at the end of each quarter. While these assessments helped us know whether students were learning, they didn’t tell us soon enough to intervene and didn’t provide specific enough information about what students hadn’t learned. So we reviewed the research and started with Black and Wiliam’s article (1998), which only had a brief mention of frequency: “It is better to have frequent short tests than infrequent long ones. Any new learning should first be testing within a week of a first encounter, but more frequent tests are counterproductive.” Wiliam (2007) updates and clarifies the issue of frequency:

*The kinds of formative assessment practices that profoundly impact student achievement cannot wait until the end of a marking period, or even the end of an instructional unit. If students have left the classroom before teachers have made adjustment to their teaching on the basis of what they have learned about students’ achievement, then they are already playing catch up.* (p. 191)

He then refers to specific instructional practices described as “classroom assessment: minute by minute, day by day” (Leahy, Lyon, Thompson, & Wiliam, 2005); this suggests that classroom teachers will see benefits increase with daily formative practices. Similarly, Marzano (2007) examines a meta-analysis of the research on how the frequency of formative assessment affects student achievement (Bangert-Drowns, Kulik, & Kulik, 1991), specifically what happens to student achievement in a fifteen-week course of study as the frequency of assessment during that time changes. He finds that if students are given five assessments during that time, they should expect an increase of 20 percentile points. The achievement increases as the number of assessments increases.

So, how frequent is really frequent enough? The final answer probably relies most on how teachers plan to respond to students who aren’t learning and may change as teachers become more proficient at using the results of the assessment they give. At Woodlawn, frequency of assessment increased from four times a year to common formative assessments every few weeks and then to formative classroom assessments written by individual teachers as often as needed.

To facilitate decisions around frequency, principals should consider asking questions that help teachers clarify their own work: “Do you have the information you need to adjust your instruction? Do you know which students are struggling early enough to help them before the unit is completed? When students are struggling, do you know what is unclear to them? Do you have the information to determine whether students already know the information you are teaching?” These questions will help teachers decide if they need more frequent assessments.

**Won’t Additional Testing Take Time Away from Teaching?**

During initial implementation, if this question isn’t being asked publicly, principals should be aware that it is likely being asked privately. Until teachers see the benefit for their students in their classroom, they are reluctant to see the worth of more assessment. This question is most often asked when teachers have been focused on using summative tests; they see assessment as something that occurs after teaching or something that is used to identify students who haven’t learned the material. Once teachers become more familiar with formative assessment practices and how to involve students in their own learning, they see these practices as part of their instruction rather than as a separate activity.

Buhle & Blachowicz (2008/2009) suggest coaching can be an effective staff development practice during this transition to a system that uses more formative assessment. Coaches help teachers realize that formative assessment can become a part of their instructional practices.

**How Do We Make Sure We Can Write Good Assessments?**

Woodlawn teachers were also concerned about whether the assessments they were writing would provide usable information. With the myriad of test banks stocked with items already written and often linked to state standards, would it be more expeditious to use these test banks rather than write our own assessments from scratch? In an effort to get things started, principals may be inclined to look for outside resources to help their teachers, but there are other matters to consider. For example, writing and designing assessments creates embedded
opportunities for teacher learning. When a teacher designs an assessment to provide feedback about student learning, she must determine what she wants to assess, what proficiency will be for each target, and how she will respond when students didn’t learn the essential concepts. Teachers who work together in collaborative teams discuss exactly what they believe a standard means, exactly what they think students should be able to do as a result of their learning, and exactly how it will “look” when students have learned. Choosing items from a test bank eliminates all of these rich discussions.

As Carol Tomlinson (2007/2008, p. 11) tells us, “Informative assessment isn’t separate from the curriculum.” No matter how closely linked to the curriculum any external set of assessments is, it will never match the taught curriculum. This doesn’t mean that teachers can’t or shouldn’t look at the assessment materials that are part of their adopted curriculum for ideas or even sample items. But it is critical that teachers choose and edit items with care to give them information to identify students’ strengths and weaknesses. For example, elementary teachers might use the assessment materials from their reading series to find grade-level appropriate texts to use to write their common assessment. They might even use some of the questions that came with the series, but they should supplement those with their own questions to create a well-designed assessment of whether students can use the specific reading strategy taught that week.

As a principal who faced this question, I needed to consider why teachers were concerned about developing their own assessments. Generally, two areas concern teachers. Some teachers feel they don’t know how to write good test items, and many teachers are overwhelmed at the thought of all the assessments they will have to create to get the data they want. Thus, the first step for the principal is to help design a plan for implementation so that teachers aren’t overwhelmed. Teachers should start with one subject area. For example, in elementary schools, teachers start with either reading or math and get more comfortable with the process. Once they’ve had some experience working as grade-level teams in discussing learning targets and proficiency standards, they could work more independently to finish the final assessments. At middle school and high school levels, teachers who teach more than one level of a subject or more than one subject may find that completing one subject area first gives them enough experience so that when they start the second area, they complete the task more quickly.

For teachers who are concerned about how to write quality items, a review of the research should help them feel more comfortable. Classroom Assessment for Student Learning (Stiggins et al., 2004) provides a thorough discussion on the variety of assessment methods and what quality assessments look like. Sometimes teachers worry that they don’t know as much about statistics as the authors of their achievement tests and state tests. Douglas Reeves (2007) addresses the issue of “psychometric perfection” when he suggests that short, teacher-created assessments have a clear advantage for teachers who want to respond immediately to the students who need more time and support. Though short tests or quizzes administered to get immediate feedback might not be statistically perfect, the advantage of immediate feedback to teachers and students is worth the possible downside. In the worst case, Reeves suggests, a student may get additional instruction he might not need, but this situation is better than the possibility that a needy student might not get help in a timely fashion if teachers wait to design perfect tests. Teachers can design effective assessments once they understand the basics of assessment design. Principals should then facilitate discussion around assessment results and whether those results are helping teachers know how to respond in their classrooms.

**Trusting Each Other with Assessment Results**
No matter what kind of assessment system a school designs, there will be some assessments that all teachers who teach the same grade level or subject area give together. If teachers are designing common formative assessments, this will occur on a regular basis. If each teacher designs his own assessments, teachers will still have common test data from state tests. Principals who want to use data effectively must develop a culture that supports sharing of both data and instructional strategies. At my school, teachers were initially reluctant to compare their results, but over time, this issue seemed to disappear as teachers began to trust each other and to believe that sharing results made them better able to help students.

**Focus on Finding Solutions**
Principals should understand that when one teacher’s results are worse than others, two things can happen. That teacher might become concerned that she will be judged as a poor teacher—or she might ask her colleagues to
share teaching strategies so that everyone’s results improve. Principals are responsible for helping develop and support a culture that encourages the second scenario.

I used specific strategies to develop this culture. The first was to model an acceptance of data as factual information that doesn’t have to lead to making excuses. In the past, if some scores were lower than we hoped, I would have given a list of “possibilities” for why we had failed. After learning more about assessment practices, however, I guided our team to list our hypotheses for how we could improve what we were doing. We talked about the things we could control and put aside excuses based on what we couldn’t control. Sure, we wished that all of our students were ready to learn and had parents who were following through at home, but we really couldn’t control those things. Rather, we brainstormed ways to help students who weren’t supervised after school and weren’t getting their homework done.

If the only tool teachers have to respond to students who aren’t learning is to help those students in their own classrooms or keep them after school, teachers will have a more difficult time embracing frequent assessments as the solution to students not learning. If, on the other hand, the entire staff works to build a schedule and system that provides support, teachers will see assessments as a means to identifying students who will get extra help from the entire team.

**Avoid Comparisons**

Principals must be especially careful never to compare one teacher to another. While it is natural to want to use data to evaluate teachers, it is important to use this information cautiously. Encourage teacher teams to analyze their own data and create their own understanding of the strengths and weaknesses that emerge. For some teachers, this may be the first time they see their results compared to their colleagues. At my school, in the beginning, we went out of our way to make the data anonymous. Each teacher got back his or her results and the total group’s results for comparison. However, it wasn’t long before teachers asked to receive data back by teacher. Today, teachers receive reports that identify each teacher’s scores and the total team’s results. Once teachers trust each other, the information they can get this way becomes much more important than anonymity.

While state and achievement test data have been available to teachers for years, they aren’t always provided in a timely way or disaggregated by teacher. Teachers need a chance to look at their individual results, compare their results with others, and come to their own conclusions—without consequence. At Woodlawn, they certainly got better at doing so with some practice. Many teachers were surprised to discover that their colleagues used different instructional strategies than they did and that sometimes the instructional strategies changed the results.

My most important role during this time was to model how to analyze and use the data we were getting. I attended data meetings, participated in finding solutions to problems, and encouraged teachers to take risks for improvement, hoping that doing so would support a culture that valued results. Principals should ensure that teachers are comfortable understanding their results and how to use them in their classrooms to help students.

**Managing Data**

If teachers don’t get their results back in a format that allows them to understand the information or quickly enough to respond the next day in their classroom, they are unlikely to effectively use the assessment data to benefit their students. Thus, principals will want to consider what type of data management system they will need to accelerate data analysis for teachers.

Data management most likely will require technological support. In my own school, this meant that I had to be willing to pilot a new computerized system of data analysis. Even though I had more tasks already on my “to do” list than I really felt I could accomplish, in the long run, putting in the time to learn and modify the program was well worth it. I learned so much about how the programs worked and what they could do that I was able to help teachers through the initial “growing pains” of implementation. It is important for principals to familiarize themselves with the technology their teachers are using, to make it as user-friendly as possible. No matter what your own comfort level with technology is, you should be willing to “feel the teacher’s pain”!
So, what did the data management system do that was helpful? First of all, teachers need to be able to use selected-response (multiple choice, matching, and true/false) questions as well as rubric data. The data management system should be able to accommodate both data types. The data system must also be able to aggregate data from students to compile classroom, team, and school results; that is, teachers should be able to look at an individual student’s results, the total results for a whole class, and (for middle school and high school teachers) all of their classes together. If teachers are writing common assessments, they need to see their data compared to the other teachers on their team. Equally important, the system should allow teachers to disaggregate their data in various ways, such as by learning target or standard and by individual students. My teachers particularly liked the reports that listed which students hadn’t mastered specific learning targets.

Teachers will also benefit from item analysis data. This allows a teacher to examine specific questions to find out what students knew and didn’t know. For example, in a multiple-choice question, a teacher might add certain distractors to see which ones students pick when they don’t know the correct answer. Then the teacher has a better understanding about how to respond to students who didn’t master the material. Teachers will also want to be able to link specific questions to state or local learning standards so that they can monitor progress around each standard.

Most of all, the system should allow teachers to administer a selected-response assessment and receive the data back in very little time. If the purpose of the assessment is to know how to respond immediately to a student who hasn’t learned a concept, the teacher really needs to know by the next day which students need help. Principals can work out ways to simplify the process of gathering and analyzing data. For example, in my school, we were able to train a very capable instructional assistant to run the computer system and print the reports. Teachers could anticipate getting their results the next day, in plenty of time to react to student needs. For extended response and performance assessments, of course, the time it takes for the teacher to evaluate student work and provide feedback will necessarily delay having the results immediately. However, teachers can use these systems to analyze results to see patterns related to learning for their students and the students on their team.

Is there a point at which too much information gets in the way? The system we used provided teachers with some statistical information that was more than they needed or wanted to have. Rather than trying to make all teachers relearn the dreaded statistics course they had to take in college, we agreed that we would leave the information on the reports in case they became useful in the future, but that teachers would only use the data they found helpful.

Using Data to Improve Student Achievement
The real power of assessment data lies in how teachers respond to it in their classrooms. If teachers fail to respond or just reteach using the same methods they used the first time, there is little likelihood that needy students will benefit. If, on the other hand, teachers use the results to identify which students need more help, and more specifically, exactly what kind of help they need, there is a strong expectation that needy students will benefit.

Thomas Guskey (2007/2008) calls these responses “corrective activities” and suggests that teachers must use different strategies that meet the specific learning needs of each student. He recommends that teachers share their instructional practices so that they each have multiple ways to teach a concept. If teachers spend time early in a unit making sure that students understand initial content, Guskey argues, they will be able to make up that time once students are confident learners.

Principals must be explicit about the need to respond to the data. Facilitate discussions about how to respond, both individually in classrooms and collectively through team and school intervention programs. As teachers become comfortable with the fact that something will happen for students as a result of assessment data, they become more comfortable with developing and using additional assessments.

Looking Forward
Once teachers become familiar with writing and using assessments, a new concern will likely emerge: teachers who are using formative assessments in their classrooms or on their teams to guide their instructional practices will want to rethink how they are using these assessments in determining student grades. Stiggins et al. (2004) suggest that
when using assessment for learning “the grading function is laid aside. This is not about accountability—those are assessments OF learning. This is about getting better” (p. 31).

Traditional grading and reporting systems are based on an average of all of the student’s work during the reporting period. Using formative assessments as feedback about what has been learned and what is still unclear assumes that learning isn’t static, but that students will continue to learn more as time goes on (O’Connor, 2002). In fact, Jacqueline Clymer and Dylan Wiliam (2006/2007) state, “The second requirement of an assessment system that supports learning is that it should be dynamic rather than static. Grades based on the accumulation of points over time are counter-productive. “They suggest that a grade should represent learning at the end of the grading period. Thomas Guskey and Jane Bailey (2001) address some teachers’ concern that students might not take formative assessments seriously by suggesting that counting these assessments for grades will limit students’ creativity and willingness to take risks as they learn new material. Rather, they advise teachers to make sure students understand the purpose of these assessments.

These questions are also likely to cause teachers to address other issues around grading practices: what to do with late work, should zeroes be permitted, should each teacher use his or her own grading system, or should all teachers use a schoolwide system, and so on. Principals must be prepared to address these questions by reading widely about reporting and grading practices and encouraging teachers to do the same.

**Student Learning Is Our Goal!**

If student learning is our goal, we must also be willing to continuously learn ourselves. When my teachers began addressing the issue of assessment, I didn’t feel prepared to answer their questions and concerns. I wanted that one guidebook that had all the answers. What I discovered was that as a profession, we are still learning about good assessment practices. As educators, we must be open to changing our thinking as we learn more through professional learning and feedback in our own classrooms and schools.

**References**


