Taking Formative

When an urban high school embraced formative assessments, teaching moved from well-intentioned guesswork to a finely-tuned dance.

Douglas Fisher, Maria Grant, Nancy Frey, and Christine Johnson

In recent years, educators have experienced much outside pressure to raise student achievement. To avoid falling into reactive and sometimes prescriptive teaching with prepackaged lessons, teachers and schools must increase the precision of our teaching (Fullan, Hill, & Crévola, 2006). This is where formative assessment comes in. Formative assessment strategies—such as oral questioning, writing prompts, and tests (Fisher & Frey, 2007)—are essential if we are to develop the detailed knowledge of students’ understandings and misunderstandings necessary to teach with precision.

Although educators have learned a lot about good formative assessment in individual classrooms, we wondered what might happen if a school took the process schoolwide. In 2001, through a joint project between San Diego State University and the San Diego Unified School District, we set out to answer this question. Hoover High School, an urban school of 2,300 students in San Diego, California, with a high percentage of low-income students and English language learners, was our test site.

The school arranged and paid for two professors from San Diego State University (coauthors Douglas Fisher and Nancy Frey) to teach part-time at Hoover High for two years while collaborating with teachers to embed a formative assessment approach in the school culture. Teachers refined a process for looking at student assessments collaboratively and using the information gathered to guide their instruction. In creating this process, we didn’t want teachers to simply give more assessments; we wanted them to see an immediate value in the process. We learned a lot from the work of Georgea Langer and her colleagues (Langer, Colton, & Goff, 2003). But rather than offering common formative assessments as an option, as Langer’s group did, we made developing and using common assessments an expectation schoolwide.

We offer here the four-step process we created for powerful schoolwide assessments at Hoover High as a model for others considering this approach.

**Step 1: Developing Pacing Guides**

Essential to this schoolwide process is the weekly meeting of teachers in course-alike groups rather than departments (for example, all teachers teaching Algebra I or world history). As a beginning point, course-alike groups developed common pacing guides. Pacing guides generally identify when the teacher will teach specific content standards, which instructional materials are appropriate, and what types of instructional strategies teachers can deploy. In addition to identifying these components, Hoover’s pacing guides also indicate key vocabulary students will need to master in order to grasp course content, which formative and summative assessments teachers will use to determine student understanding, and what accommodations are recommended for students with disabilities, English language learners, or students performing above grade level.

64 EDUCATIONAL LEADERSHIP/DECEMBER 2007/JANUARY 2008
Assessment Schoolwide

Step 2: Designing Common Assessments
In addition to choosing pacing guides and corresponding summative assessments, teacher groups at Hoover design, develop, or modify assessment items that every teacher will administer regularly throughout that course. Teachers develop these test items in such a way as to provide information that will help them determine what content students understand, where students have gaps in comprehension, and who needs intervention. As groups of teachers develop these assessment items, they learn more about their state's content standards and how those standards might be assessed on state tests. In addition, they plan items that will signal when students are overgeneralizing, oversimplifying, or exhibiting common misunderstandings. We learned about assessment design along the way, learning from assessments we wrote that didn't work and from professional development seminars we attended.

As part of designing common assessments to use throughout the year, Hoover teachers generally create some common formative assessment items that mirror the state test design because they know that test format practice is essential. Students must understand tests as a genre—how they work and what to expect. However, teachers do not limit items to those that mirror the state test format: They also include short-answer, constructed-response, and alternative-response items, as well as timed essays. We know that it's best to rely on a number of strategies for determining students' understanding and that the key to taking formative assessments schoolwide is ensuring that teachers can determine "next steps" in instruction on the basis of such assessments—which requires more than practicing standardized questions.

Step 3: Conducting Item Analysis
Teachers in course-alike groups engage in the third step, analyzing the results, after all students in that course have participated in a common formative assessment. At Hoover, teachers use Edusoft, one of several commercial software programs that provide an item analysis for each assessment and indicate the percentage of students who selected each of the answers. Other powerful programs include Datawise and Instructional Data Management System.

To avoid falling into prescriptive teaching, teachers and schools must increase the precision of their teaching.

The item analysis is key to instructional conversations and the interventions that flow from them because it enables teachers to look across the student body for trends—content or concepts they need to reteach, assessment items they need to change, or pacing guides they need to revise. Edusoft also enables teachers to analyze the results of clusters of students, such as exploring how English language learners as a group performed on a specific item.
We didn’t want teachers to simply give more assessments; we wanted them to see an immediate value in the process.

**Step 4: Engaging in Instructional Conversation**

The fourth step, instructional conversation, is why Hoover High teachers do all this work. Talking with colleagues who teach the same content and see the same data results is foundational to instituting improvements and helps teachers determine which instructional strategies are working, which materials are effective, and which students still need help to master the standards. Each course group has a leader who receives professional development in facilitation skills. Such conversations enable teachers to return to their individual classrooms and engage in the real work of formative assessments—to reteach and intervene where students aren’t doing well.

Let’s consider two fruitful instructional conversations observed recently at a regular weekly meeting at Hoover.

**Uncovering Gaps in Genetics Knowledge**

Hoover science teacher Maria Grant regularly facilitates conversations about student work. She and her colleagues teaching 10th grade biology recently had the following conversation about students’ understandings of genetics concepts while examining students’ responses to a common formative assessment:

In a certain species of insect, the allele for brown eyes (B) is dominant to the allele for blue eyes (b). For this species, eye color does not depend on the sex of the organism. When a team of scientists crossed a male and a female that both had brown eyes, they found that 31 offspring had brown eyes and 9 had blue eyes. What are the most likely genotypes of the parent insects?

A. BB and bb
B. bb and bb
C. Bb and Bb
D. BB and Bb

Each answer shows the two alleles for eye color of the male and female insect. The correct answer, which 46 percent of the students chose, is C because most of the offspring have brown eyes but a few have blue eyes. For an offspring to have blue eyes, it must receive a b allele from both parents (bb); a combination of Bb or BB would yield brown eyes.

Mr. Simms began the discussion:

**MR. SIMMS:** The greatest percentage of students did choose the correct answer.

Ms. Jackson: Yes, but 54 percent didn’t choose the right answer. 17 percent chose answer A. This might mean that students don’t understand how a recessive trait is passed on.

**MR. SIMMS:** Even though I covered the main concepts of Mendelian genetics, it seems that students didn’t really understand how expressed traits are passed from parent to offspring.

**MRS. RODRIGUEZ:** Yes, and 11 percent chose answer B. The students that chose this answer don’t understand the concept of a dominant allele. Maybe I need to focus more on vocabulary instruction for this group of students. We covered the key terms, but they don’t seem to know how to use them. I think we should find out the specific students who missed this and get to them during small group time.

Ms. Jackson: I also think we need to work on test-taking skills. Our students should have been able to eliminate answers A and B right away because each shows a parent with blue eyes, and the question states that both parents have brown eyes.

**MR. SIMMS:** Twenty-six percent of students chose answer D. Maybe they thought that since three out of four alleles are B, there’s a correlation to the 31 out of 40 total offspring with brown eyes described in the question. I think I need to review how to use Punnett’s squares.

Ms. Grant: Maybe if we shared these results with students, it would facilitate their thinking about the content. What if we showed all students this item analysis and asked them to work in small groups to determine why specific answers were wrong?
Wouldn't that help them get test-format practice and also reinforce the biology?

By the end of this conversation, the teachers decided to reteach some basic concepts and show the students the item analysis to focus them on the reasons for the correct and incorrect answers.

**Parsing Mastery of a History Unit**
Hoover's history teachers also analyze common formative assessments and change their teaching strategies on the basis of what they find. The department recently piloted a metacognitive task in combination with a content assessment. For each question students answered, they also indicated one of the following four descriptions of how they answered: I knew it, I figured it out, I guessed at it, or I don't care. During a discussion of this assessment, for the 9th grade course Foundations of Democracy, teachers examined a question that confused a number of students (see fig. 1). Mr. Jacobs summarized the knowledge gaps this question showed:

Let's start with Question 3. Only 61 percent of the students got it right, and only 38 percent of those who answered correctly self-reported that they knew it. An additional 36 percent said they figured it out, and 24 percent guessed at it. It's interesting that only 3 kids (of 241) didn't care about this question. I know that I taught this. Most of the wrong answers were still based on [students' understanding of] democracy, but not the right type of democracy. I think this could be a quick fix. We need to make sure that students really have a sense of the difference between direct and representative democracy. I have an idea for a simulation that could solidify this for students.

Mr. Jacobs described his idea for a simulation, and the teachers agreed to reteach this concept. Mrs. Johnson then turned their attention to Question 10:

**FIGURE 2. Changes in Student Achievement at Hoover High**

<table>
<thead>
<tr>
<th>How Students Scored in Biology</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Proficient</td>
<td>1%</td>
<td>18%</td>
</tr>
<tr>
<td>Basic</td>
<td>28%</td>
<td>51%</td>
</tr>
<tr>
<td>Below basic</td>
<td>42%</td>
<td>22%</td>
</tr>
<tr>
<td>Far below basic</td>
<td>27%</td>
<td>8%</td>
</tr>
<tr>
<td>(n = 333)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How Students Scored in History</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>0%</td>
<td>3%</td>
</tr>
<tr>
<td>Proficient</td>
<td>3%</td>
<td>13%</td>
</tr>
<tr>
<td>Basic</td>
<td>27%</td>
<td>26%</td>
</tr>
<tr>
<td>Below basic</td>
<td>24%</td>
<td>27%</td>
</tr>
<tr>
<td>Far below basic</td>
<td>46%</td>
<td>31%</td>
</tr>
<tr>
<td>(n = 563)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data reflect 10th grade scores on the California Standards Test before and after Hoover High implemented schoolwide formative assessment.

Ms. Vasquez confessed, "I don't really know how to teach this. I've shown my students the map and the directions. I don't know what to do differently." Mrs. Johnson suggested to her, "I'll cover your class so that you can go watch Mr. Applegate teach this concept. Is that OK?" She then asked, "Does anyone else need help with teaching cardinal points?" Because many teachers wanted help, Mrs. Johnson recommended that the group consider revising the course's pacing guide to allow more time to teach map skills.

As they continued to analyze the results, the teachers also identified a small group of students who had missed all the test items related to government structures. They believed these learners would benefit from instruction to build their background knowledge of such structures. Mr. Applegate offered to meet with these students during the school's after-school tutoring time.

Teachers also examined the students' self-assessments and determined a correlation between accuracy and a response of "I knew it." Students who checked the "I figured it out" indicator were also often accurate. The teachers were pleased to see students using test-taking strategies of elimination and using context clues.

**The Fruits of Precision Teaching**
Although the joint action project with San Diego State has ended, Hoover teachers continue to engage in some step of this four-step process every...
week, on the day students are released early. Hoover has experienced impressive gains in student achievement since adopting formative assessment schoolwide. As data shown in Figure 2 indicate, average student performance on the California Standards Test in both biology and history improved appreciably over the first two years that Hoover High has been involved in this formative assessment process. In 2005, for example, 51 percent of Hoover's 10th graders scored at the basic level on the California Standards Test in biology, compared with only 28 percent achieving at the basic level in 2003. Similarly, in 2005, 18 percent scored at the proficient level on this test, compared with only 1 percent scoring at the proficient level in 2003.

These changes came about because all Hoover's teachers became more precise in their teaching. Collaborative item analyses and rich instructional conversations based on these analyses, characterized by collegiality and respect, drove these changes. The key to powerful formative assessment, whether schoolwide or class-specific, is for teachers to take action as soon as they have information about what students do and don't understand. With this key, we can all teach with precision. 

References

Douglas Fisher (dfisher@mail.sdsu.edu) is Professor of Literacy and Nancy Frey (nfrey@mail.sdsu.edu) is Associate Professor of Literacy at San Diego State University in California. Maria Grant (mgrant@fullerton.edu) is currently Assistant Professor of Secondary Education at California State University in Fullerton; Christine Johnson (cjohnson@hshmc.org) is an educational consultant. Fisher and Grant are authors of Better Learning Through Structured Teaching: A Framework for Gradual Release of Responsibility (ASCD, in press).