“Dylan Wiliam’s wonderful new book strips away the murk currently surrounding formative assessment. Deftly moving between research evidence and classroom practicalities, Wiliam provides a compelling vision of how formative assessment can enhance students’ learning. For today’s educators, this is truly a must-read!”

—W. James Popham, Professor Emeritus, UCLA

“With precision, clarity, and practical comprehensiveness, Wiliam gives us all we need to know to get deep and lasting success. Each of his five strategies is a gem; in combination, they are unbeatable.”

—Michael Fullan, Professor Emeritus, University of Toronto

If we want our students to thrive in the complex, unpredictable world of the 21st century, we must concentrate on increasing educational achievement by increasing the quality of the teachers in our schools. *Embedded Formative Assessment* faces this challenge head-on by making a case for the important role of formative assessment as a process, not a tool, in increasing teacher quality and thus student learning.

Author Dylan Wiliam:

• Provides over fifty practical techniques for classroom formative assessment
• Emphasizes the instructional side of the process
• Explores in-depth the use of classroom questioning, learning intentions and success criteria, feedback, collaborative and cooperative learning, and self-regulated learning to engineer effective learning environments

“Wiliam’s most recent work confirms his place in the front rank of global thought leaders in education. Teachers and school leaders will find this work enormously practical.”

—Douglas Reeves, The Leadership and Learning Center

“This is a must-have, must-read, must-put-into-immediate-action book for every teacher, principal, and network leader. Dylan has made powerful coaching learning ideas accessible in a highly engaging way and makes the development of a learning system possible.”

—Linda Kaser, Co-Director, Network of Performance Based Schools, and Co-Director, Center for Innovative Educational Leadership, VIU

solution-tree.com
Acknowledgments

This book is the result of a thirty-five year journey that has taken me from public school teacher to teacher educator to researcher and university administrator. Given the length of this journey, you will not be surprised that I have a lot of people to thank. First, the students at Christopher Wren School and North Westminster Community School in London taught me most of what I know about learning. The research I have read and carried out since then has done little more than help me make sense of what I learned there. Second are the teachers I have worked with in those two schools and in the extended networks of teachers of which I have been a part. As someone who went straight from a bachelor's degree to the classroom, with no training or preparation, I was fortunate to start teaching in one of London's toughest schools—fortunate because every teacher in that school struggled and was willing to talk about their struggles. It is only subsequently that I have realized how lucky I was, because that openness is not found everywhere.

Mentioning individuals is always invidious, but I owe particular debts of gratitude to Sarah Sharkey and Jerry Hardy (Christopher Wren School) and Dick Gmiterek (North Westminster Community School). When I joined the University of London, I lucked out again, being mentored by two great thinkers about education—first, Margaret Brown and later, Paul Black. As well as being extraordinary academics, they are two of the nicest people one could ever hope to meet, and their generosity to younger academics is a model I have tried to emulate. I am also grateful to the editorial staff at Solution Tree, and particularly Lesley Bolton, for all the work they have done in getting my manuscript to the finished book you have before you. Finally, and most importantly, my greatest thanks go to my partner, Siobhán Leahy. We first met at a mathematics teachers' conference over thirty years ago, and for thirty years, she has been my greatest inspiration, my greatest support, and has kept me grounded in the realities of daily school life. That is why the book is dedicated to her.
Solution Tree Press would like to thank the following reviewers:

Travis Bracht
Director of Student Learning
Francis Howell School District
St. Charles, Missouri

Bernice Burroughs
Principal
The St. Johnsbury School
St. Johnsbury, Vermont

Peggy Clark
Educational Consultant
The Schultz Center for Teaching and Leadership
Jacksonville, Florida

Eileen Depka
Assistant Superintendent of Educational Services
School District of Elmbrook
Brookfield, Wisconsin

Roger Jenni
Testing and Assessment Coordinator
Northfield Public Schools ISD 659
Northfield, Minnesota

Sheila Larson
Curriculum Director
Fowlerville Community Schools
Fowlerville, Michigan

Kimberly Young
Consultant, Office of Educational Assessment and Accountability
Michigan Department of Education
Lansing, Michigan
Table of Contents

About the Author ........................................................................................................... ix
Introduction ......................................................................................................................... 1

CHAPTER 1
Why Educational Achievement Matters ............................................................................. 3
   The Increasing Importance of Educational Achievement ............................................. 3
   Why Is Raising Student Achievement So Hard? ......................................................... 10
   Three Generations of School Effectiveness Research .................................................... 15
   The Impact of Teacher Quality ....................................................................................... 17
   How Do We Increase Teacher Quality? ......................................................................... 22
   Conclusion ....................................................................................................................... 26

CHAPTER 2
The Case for Formative Assessment ................................................................................... 27
   Professional Development .............................................................................................. 27
   The Origins of Formative Assessment ............................................................................ 33
   What, Exactly, Is Formative Assessment? ...................................................................... 37
   Strategies of Formative Assessment .............................................................................. 45
   Assessment: The Bridge Between Teaching and Learning ............................................. 46
   Conclusion ....................................................................................................................... 50

CHAPTER 3
Clarifying, Sharing, and Understanding Learning Intentions and Success Criteria ............... 51
   Why Learning Intentions Are Important ....................................................................... 51
   When Are Learning Intentions Useful? ........................................................................... 56
   Issues in Constructing Learning Intentions ................................................................. 61
   Practical Techniques ....................................................................................................... 65
   Conclusion ....................................................................................................................... 69

CHAPTER 4
Eliciting Evidence of Learners’ Achievement ................................................................... 71
   Finding Out What Students Know .................................................................................. 71
About the Author

Dylan Wiliam, PhD, is a consultant who works with educators in North America, the United Kingdom, and many other countries to develop effective, research-based formative assessment practices. He is former deputy director of the Institute of Education at the University of London. From teaching in urban public schools to directing a large-scale testing program to serving in university administration, his professional path has led to a variety of positions at the forefront of education. Dr. Wiliam was also, from 2003 to 2006, senior research director at the Educational Testing Service in Princeton, New Jersey.

During his early years of teaching in inner-city classrooms, Dr. Wiliam focused on physics and mathematics. He later joined the faculty of Chelsea College, University of London, which later became part of King’s College London. Here, he worked on developing innovative assessment schemes in mathematics before accepting leadership of the King’s College Mathematics Education Program.

For three years, Dr. Wiliam served as the academic coordinator of the Consortium for Assessment and Testing in Schools, which developed a variety of assessments for the national curriculum of England and Wales. He then returned to King’s College to serve as dean of the School of Education before being promoted to assistant principal of the college.

In 1998, he coauthored a major review of research evidence on formative assessment with Paul Black and has worked with many teachers in the United Kingdom and United States on developing formative assessment practices to support learning.
In addition to a doctor of education, Dr. Wiliam holds numerous degrees in mathematics and mathematics education. To learn more about Dr. Wiliam’s work, visit www.dylanwiliam.net.

To book Dr. Wiliam for professional development, contact pd@solution-tree.com.
Introduction

In 1984, I left my job teaching math and science in an urban public school in London to join a research project at the University of London that was exploring the potential of formative assessment to improve student learning. Over a quarter of a century later, this book is the result of that journey.

The book has two main purposes. The first is to provide simple, practical ideas about changes that every teacher can make in the classroom to develop his or her practice of teaching. The second is to provide the evidence that these changes will result in improved outcomes for learners.

In chapter 1, I show why educational achievement is so important and why raising educational achievement needs to be a national priority; with higher levels of educational achievement, people are healthier, live longer, contribute more to society, and earn more money. For society, the benefits include reduced criminal justice costs, reduced health care costs, and increased economic growth.

I also outline briefly how previous attempts at reform—including changes to the structure of schooling, to the governance of schools, and to the curriculum, and an increased role for digital technology—have been largely ineffective. These reform efforts have failed to take into account three crucial points:

1. The quality of teachers is the single most important factor in the education system.
2. Teacher quality is highly variable.
3. Teacher quality has a greater impact on some students than others.

In chapter 1, I also show that attempts at improving the quality of entrants into teaching and removing the least effective teachers will result in small effects that will take a generation to materialize. The fundamental argument of chapter 1, therefore, is that to secure our future
economic prosperity, we need to help the teachers who are already serving in our schools improve.

In chapter 2, I explore some of the ways that teachers might develop and estimate how big an impact these changes would have on student outcomes. I show that some popular initiatives, such as learning styles, have no discernible impact on student achievement at all, while others, such as increasing teacher content knowledge, do improve student learning but by much less than is generally assumed.

In the second part of chapter 2, I summarize the research on classroom formative assessment practices and show that these practices appear to have a much greater impact on educational achievement than most other reforms. Chapter 2 concludes by outlining what formative assessment is, and what it is not, and presents the five key strategies of formative assessment.

Chapters 3, 4, 5, 6, and 7 deal in turn with each of the five key strategies of formative assessment:

1. Clarifying, sharing, and understanding learning intentions and criteria for success
2. Engineering effective classroom discussions, activities, and learning tasks that elicit evidence of learning
3. Providing feedback that moves learning forward
4. Activating learners as instructional resources for one another
5. Activating learners as owners of their own learning

In each of these five chapters, I present a summary of the research evidence that shows the impact of the strategy, and I offer a number of practical techniques that teachers have used to incorporate the strategy into their regular classroom practice. Although there is a definite progression throughout the chapters, I have made each chapter as self-contained as possible.

In all, these five chapters describe over fifty practical techniques for classroom formative assessment. Most of these techniques are not new; what is new is the framework for formative assessment that is presented in chapter 2, which shows how these disparate techniques fit together, and the research evidence that shows that these are powerful ways of increasing student engagement and helping teachers make their teaching more responsive to their students’ needs.
Providing Feedback That Moves Learning Forward

It seems obvious that feedback to students about their work should help them learn, but it turns out that providing effective feedback is far more difficult than it appears. Much of the feedback that students get has little or no effect on their learning, and some kinds of feedback are actually counterproductive. This chapter reviews the research on feedback; why some kinds of feedback are, at best, useless and, at worst, actually lower performance; and how teachers can give their students feedback that moves learning forward.

The Quality of Feedback

The power of feedback to improve classroom learning was vividly demonstrated in a study by Maria Elawar and Lyn Corno (1985). A group of eighteen sixth-grade teachers in three schools in Venezuela received seven hours of training on how to provide constructive written feedback on the mathematics homework produced by their students. The feedback included specific comments on errors, suggestions to the students about how to improve, and at least one positive remark. A second group of teachers received the same training as the first group but gave constructive feedback to half their classes and just scores to the other half. A third group of teachers received no training and graded...
homework as normal (that is, by giving only scores). The students receiving the constructive feedback learned twice as fast as the control-group students—in other words, they learned in one week what the other students took two weeks to learn. Furthermore, in the classes given constructive feedback, the improvement in learning occurred across the achievement range, attitudes toward mathematics were more positive, and the achievement gap between male and female students was reduced (Elawar & Corno, 1985).

Ruth Butler (1988) investigated the effectiveness of different kinds of feedback on 132 sixth-grade students in twelve classes in four schools in Israel. For the first lesson, the students in each class were given a booklet containing a range of divergent thinking tasks in which students were asked to identify unusual uses for familiar objects (Torrance, 1962). At the end of the period, their work was collected. This work was then graded by researchers and teachers working independently. At the beginning of the next period, two days later, the students were given feedback on the work they had done in the first period. In the first classroom of each of the four schools, students were given scores: the lowest-quality work was given a score of 40; the highest-quality work was given a score of 99; and other work was given a score between these two extremes. In the second classroom in each of the four schools, students were given comments, such as “You thought of quite a few interesting ideas; maybe you could think of more ideas.” In the third classroom in each of the four schools, the students were given both scores and comments. Then the students were asked to attempt some similar tasks and told that they would get the same sort of feedback as they had received for the first lesson’s work. Again, the work was collected and scored.

Those given only scores made no progress from the first lesson to the second—their work was no better. When asked whether they wanted to continue doing similar work, those who had received high scores indicated that they did, but those who had received low scores did not. The students given only comments scored, on average, 30 percent higher on the work done in the second lesson than that done in the first (although, of course, they did not know this because they had not been given scores), and all these students indicated that they wanted to carry on doing similar work. The obvious question is, What happened to the students given both scores and comments?
When I ask teachers what they think happened in this study, many assume that the students given both scores and comments progressed at least as much as those given comments. After all, if the comments alone produced a 30 percent improvement in scores, then giving both scores and comments should have been even more informative.

Other teachers suggest that there was some trade-off between the two kinds of feedback. For example, some teachers say that they expect an improvement in scores because of the effects of the comments, but there may have been some polarization of attitudes as a result of the scores; in other words, students who got high scores were motivated and wanted to carry on, while low scorers were demotivated and wished to move on to something else. A second possible trade-off works the other way, with students making no progress due to the effect of the scores but with attitudes kept positive by the encouragement provided by the comments.

Most teachers, therefore, are surprised to learn that the effect of giving both scores and comments was the same as the effect of giving scores alone. Far from producing the best effects of both kinds of feedback, giving scores alongside the comments completely washed out the beneficial effects of the comments; students who got high scores didn’t need to read the comments, and students who got low scores didn’t want to. And yet giving both a score (or grade) and some kind of comment is probably the most prevalent form of feedback to students in the United States. This study (and others like it, to follow) shows that if teachers are providing careful diagnostic comments and then putting a score or a grade on the work, they are wasting their time. They might as well just give a score or a grade—the students won’t learn anything as a result, but the teacher will save a great deal of time.

Some clues as to why feedback can have these effects are provided by another study by the same researcher (Butler, 1987). This time, 200 fifth- and sixth-grade students in eight classes spent a period working on a variety of divergent thinking tasks. Again, the work was collected and the students were given one of four kinds of feedback on this work at the beginning of the second period (again two days later):

1. In two classes, the students were given comments.
2. In two classes, the students were given grades.
3. In two classes, the students were given written praise.
4. In two classes, the students were given no feedback at all.
The quality of the work done in the second period was compared to that done in the first. The quality of the work of the students who had been given comments had improved substantially compared to their work in the first period, but those given grades and praise had made no more progress than those given absolutely no feedback on their work.

At the end of the second period, the students were given a questionnaire. The questionnaire sought to establish what factors influenced the students’ decisions about whether to expend effort in the classroom and what factors the students believed determined whether they were successful or not. Specifically, the questionnaire was designed to elicit whether the students attributed their expenditure of effort and their success to ego-related factors or to task-related factors, as shown in table 5.1.

<table>
<thead>
<tr>
<th>Attribution of</th>
<th>Ego</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure of effort</td>
<td>To do better than others To avoid doing worse than others</td>
<td>Interest To improve performance</td>
</tr>
<tr>
<td>Success</td>
<td>Ability Performance of others</td>
<td>Interest Effort Experience of previous learning</td>
</tr>
</tbody>
</table>

Students who were given comments had high levels of task-involvement, but their levels of ego-involvement were the same as those given no feedback. However, while those given written praise and those given grades had comparable levels of task-involvement to the control group, their levels of ego-involvement were substantially higher. As noted, the provision of grades and written praise had no effect on achievement; their only effect was to increase the sense of ego-involvement. This, as anyone involved in guidance and counseling work in schools knows, is bad news. To change a student’s behavior, it is generally far more effective to focus on criticizing the behavior rather than the student (in effect, giving task-involving rather than ego-involving feedback).

These findings are also consistent with research carried out in the 1970s that showed clearly that praise was not necessarily a good thing—in fact, the best teachers appear to praise slightly less than average (Good & Grouws, 1975). It is the quality rather than the quantity of praise that is important, and in particular, teacher praise is far more effective if it is
Providing Feedback That Moves Learning Forward

infrequent, credible, contingent, specific, and genuine (Brophy, 1981). It is also essential that praise is related to factors within an individual’s control, so praising a gifted student simply for being gifted is likely to lead to negative consequences in the long term (Dweck, 2006).

The timing of feedback is also crucial. If it is given too early, before students have had a chance to work on a problem, then they will learn less. In a review of forty research reports on the effects of feedback in “test-like” events—such as questions embedded in programmed learning materials, review tests at the end of a block of teaching, and so on—Robert Bangert-Drowns, Chen-Lin Kulik, James Kulik, and MaryTeresa Morgan (1991) found that what mattered was the degree of “mindfulness” in the students that the feedback generated. When students could peek ahead and look at answers to questions before they had tried to answer them, they learned significantly less than when they had to attempt to answer the question before getting feedback.

A direct demonstration of this was provided in a study by Malcolm Simmons and Peter Cope (1993). Pairs of students aged between nine and eleven worked on angle and rotation problems. Some students worked on the problems on a computer, using the programming language Logo, and some worked on the problems using pencil and paper. The students working in Logo were able to use a “trial and improvement” strategy that enabled them to get a solution with little mental effort. For those using pencil and paper, working out the effect of a single rotation was much more time consuming, giving these students an incentive to think carefully, and this greater “mindfulness” led to more learning.

The key idea in all this—that what matters is the mindfulness with which students engage in the feedback—means that sometimes less is more. A study of sixty-four third graders required the students to engage in arithmetical reasoning tasks and varied the kind of support they received. Half of the students were given a scaffolded response when they got stuck—in other words, they were given only the minimum amount of support to get them unstuck and to make progress. The other half of the students were given a complete solution to the problem on which they were stuck and then given a new problem to work on. Students given the scaffolded response learned more and retained their learning longer than those given full solutions (Day & Cordón, 1993). In a sense, this is hardly surprising, since those given the complete solutions
had the opportunity for learning taken away from them. As well as saving the teachers time, developing skills of intervening as little as possible—just enough to get the students going—promotes better learning. A good example of this kind of feedback is provided by Jonathon Saphier (2005):

Teacher: “What part don’t you understand?”

Student: “I just don’t get it.”

Then we add one more variable. See if you can find out what it is, and I’ll come back in a few minutes.” (p. 92)

Most teachers have had the experience of giving a student a new task only for the student to ask for help immediately. When the teacher asks, “What can’t you do?” a common reply is, “I can’t do any of it.” In such circumstances, the student’s reaction may be caused by anxiety about the unfamiliar nature of the task, and it is often possible to support the student by saying something like, “Copy out that table, and I’ll be back in five minutes to help you fill it in.” This is usually all the support the student needs. Copying out the table forces the student to look in detail at how the table is laid out, and this busywork can provide time for the student to make sense of the task herself.

Another good example of scaffolded feedback is provided by Ian Smith (2008). An art teacher had sketched a face and, in collaboration with the class, determined seven criteria for a successful portrait, including the idea that the eyes should be halfway down the face and that the distance between the eyes should be roughly the same as the width of one eye, and so on. She was then able to give feedback to a student in the format shown in figure 5.1.

![Figure 5.1: Feedback grid for a face-drawing task.](image)

This communicated clearly to the student what needed attention, took up very little of the teacher’s time, and still left plenty of work for the student to do.
An obvious question that arises from these studies is, Does it make a difference whether the feedback is given orally or in writing? One of the few studies to address this was that of eighty Canadian students in three groups learning to write the major scales in their music classes (Boulet, Simard, & De Melo, 1990). One group was given written feedback, a list of weaknesses, and a work plan; the second group was given oral feedback on the nature of their errors, plus a chance to work on improvement in class; and the third group was given no feedback. At the beginning of the study, there were no differences between the three groups in terms of their previous achievements in music, scores on a test of musical aptitude, and on a measure of their ability to learn. All groups fell short of the mastery level of 80 percent set for this task, but the students in the second group, who had been given oral feedback, scored significantly higher than the students in the other two groups (which were not significantly different from each other). Because the treatment of group 2 differed in three ways from that of group 1, it is not possible to determine what made the difference. However, from their observations, the researchers indicated that whether the feedback was given orally or in writing was much less important than the fact that group 2 was given time, in class, to use the feedback to improve their work.

Some types of feedback actually lower performance. Avraham Kluger and Angelo DeNisi looked at every study they could find that had ever been done on the effects of feedback in schools, colleges, and workplaces over a ninety-year period (from 1905 to 1995). They defined a feedback intervention as “actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one’s task performance” (Kluger & DeNisi, 1996, p. 255) and found over 3,000 research studies (2,500 journal articles and 500 technical reports) that looked at the impact of such feedback interventions on performance.

Of course, these studies varied in their quality, and to be sure that poor-quality studies were not being included, Kluger and DeNisi established a number of criteria for inclusion in their review. First, there had to be two groups of participants for whom the only difference (as far as could be judged) was whether they had received feedback or not. Second, a study had to have had at least ten participants (some, particularly in the medical field, had only a single participant). Third, the study had to include some kind of measurement of performance with
sufficient details of the measurements of impact provided to be able to calculate the size of the impact of feedback on performance.

Astonishingly, only 131 of the original 3,000 studies—approximately 4 percent—satisfied these criteria of scientific quality. This figure was so low that Kluger and DeNisi reviewed the excluded studies to make sure that worthwhile studies were not being unduly rejected and concluded that those studies rejected could not shed useful light on the effects of feedback.

Just as surprisingly, in 50 of the 131 accepted studies, providing feedback actually lowered performance. In other words, in almost two out of every five carefully conducted studies, the participants would have done better if the feedback had not been given! To try to understand why feedback could have such counterproductive—and unexpected—effects, Kluger and DeNisi looked in detail at the studies to determine when feedback does and does not improve performance. They pointed out that when the feedback draws attention to a gap between one’s current performance and the goal, what happens depends on whether the current performance is higher or lower than the goal.

When the feedback tells an individual that he has already surpassed the goal, one of four things can happen. Obviously one hopes that the individual would seek to change the goal to one that is more demanding, but it might also be taken as a signal to ease off and exert less effort. Also, when success comes too easily, the individual may decide that the goal itself is worthless and abandon it entirely or reject the feedback as being irrelevant.

When, as is more common, the feedback indicates that current performance falls short of the goal, there are again four responses. The recipient may change the goal, for example, when a student decides to settle for a B even though an A might be within her grasp but is felt to be too much work or too risky in terms of one’s self-image to attempt. A second response is to abandon the goal altogether, as seen in the “learned helplessness” (Dweck, 1975) displayed by students who, in the face of repeated indications that they are falling short, decide that something is forever beyond them—for example, when a student says, “I’m no good at math.” The third type of response is to reject the feedback. This is quite common in workplace settings when, for example, someone who feels he has contributed well beyond what was expected gets a neutral evaluation from a supervisor. The effect of such rejection of feedback is to lower commitment (and therefore, possibly, to reduce performance). The fourth kind of response—and presumably the
Providing Feedback That Moves Learning Forward

one intended by the person giving feedback—is to change one’s behavior so
as to increase performance to match the goal. These responses to feedback
are summarized in table 5.2.

Table 5.2: Possible Responses to Feedback

<table>
<thead>
<tr>
<th>Response type</th>
<th>Feedback indicates performance exceeds goal</th>
<th>Feedback indicates performance falls short of goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change behavior</td>
<td>Exert less effort</td>
<td>Increase effort</td>
</tr>
<tr>
<td>Change goal</td>
<td>Increase aspiration</td>
<td>Reduce aspiration</td>
</tr>
<tr>
<td>Abandon goal</td>
<td>Decide goal is too easy</td>
<td>Decide goal is too hard</td>
</tr>
<tr>
<td>Reject feedback</td>
<td>Ignore feedback</td>
<td>Ignore feedback</td>
</tr>
</tbody>
</table>

Only the two italicized responses are likely to improve performance. The other six, at best, do nothing and, at worst, lower performance, sometimes a considerable degree.

The research reviewed by Kluger and DeNisi (1996) also shows that it is very difficult, if not impossible, to predict which of these responses will occur. It will depend on the individual receiving the feedback, the kind of task on which feedback is given, and the recipient’s perceptions of the person giving the feedback.

In concluding their review of the effects of feedback, Kluger and DeNisi point out that the biggest problem is the failure of most of the studies they reviewed to consider long-term effects. Even when the effect of feedback is to increase motivation, this may be a short-term improvement at the expense of a long-term worsening of performance. They suggest that, instead, research on feedback should focus less on the actual impact on performance and more on the kinds of responses that are triggered in the individual as a result of the feedback.

In a series of research studies spanning over thirty years, Carol Dweck of Stanford University and her colleagues have investigated, among other things, how students make sense of their successes and failures in school by asking students questions such as, “When you get an A, why is that?” or “If you got an F, why might that be?” Dweck and her colleagues found that there were three strong themes running through the students’ responses (Dweck, 2000).

The first was whether the success or failure was due to factors relating to the individual or due to outside factors (in other words, how the
attribution was *personalized*. For example, “I got an A because I did a good piece of work” would be an *internal* attribution because the cause is seen as being within the individual. “I got an F because Mr. Smith hates me” would be an *external* attribution, because the cause is seen as being outside the individual. It is quite common to find individuals attributing their successes internally and their failures externally. There is an old Chinese adage that captures this tendency beautifully: “Success has a thousand fathers, but failure is an orphan.”

The second theme was whether success was seen as being due to factors that were likely to be long lasting or transient (in other words, the permanence or stability of the factor). For example, when attributing success, “being smart” is likely to be seen by students as *stable* (that is, long lasting); if the student got an A this time because she is smart, she is likely to get an A next time because she will still be smart. If, on the other hand, a student says, “I got an A because I worked really hard on this assignment,” then this is attributing success to an *unstable* (that is, transient) factor. Next time, whether the student gets an A will depend on how much effort he puts into the assignment.

The third was the *specificity* of the attribution: whether success or failure is seen as being due to factors that affect performance in all areas or just the area in question. For example, some students overgeneralize their successes. They believe that because they are told they are good at one thing, they are good at everything. Students can also overgeneralize failures and believe that because they fail at one thing, they are a failure at everything. (These are both *global* attributions.) Table 5.3 gives some examples of attributions of success and failure (see also Dweck, 2006).

Dweck (2000) and others have found that there is a slight tendency for boys to be more likely to attribute their successes to stable causes (such as ability) and their failures to unstable causes (such as lack of effort and bad luck). This would certainly explain the high degree of confidence with which many boys approach examinations for which they are rather unprepared. More controversially, the same research suggests that there is a slight tendency for girls to attribute their successes to unstable causes (such as effort) and their failures to stable causes (such as lack of ability), leading to the learned helplessness mentioned previously.

It is important to remember that these are not sharp differences between boys and girls—the overlap in how males and females attribute successes is far, far greater than the differences between males and
Providing Feedback That Moves Learning Forward

females—but these differences may in part explain why female students are now outperforming male students all over the world. The best learners consistently attribute both success and failure to internal, unstable causes. They believe: “It’s up to me” (internal) and “I can do something about it” (unstable). The tendency is for girls to do this for their successes, although not for their failures, while the tendency for boys is to attribute neither successes nor failures to internal, unstable causes (successes tend to be attributed to internal stable causes, while failures tend to be attributed to external unstable causes). Regardless of the current tendencies, learning in classrooms will be considerably enhanced if students embrace this idea of “It’s up to me, and I can do something about it.” When I talk to students about these issues, I often use examples from sports, such as Michael Jordan, Tom Brady, and Mike Piazza.

Although it is not true that Michael Jordan was cut from his high school varsity basketball team, he did experience a significant reversal. His high school, Lacey High School, ran a junior varsity team in addition to the main varsity team. Although the expectation was that most tenth graders would play on the junior varsity team, some exceptionally gifted athletes in the tenth grade did play on the varsity team. Jordan and his friend Leroy Smith attended a basketball camp in the summer after ninth grade and were encouraged by the coach of the varsity team, Pop Herring, to try for the varsity team in the tenth grade. In the end, Smith made it, and Jordan did not. The assistant head coach at the time, Fred Lynch, said, “Leroy was not a better basketball player than Mike,

<table>
<thead>
<tr>
<th>Attribution</th>
<th>Success</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalization</td>
<td>Internal: “I got a good grade because it was a good piece of work.” External: “I got a good grade because the teacher likes me.”</td>
<td>Internal: “I got a low grade because it wasn’t a very good piece of work.” External: “I got a low grade because the teacher doesn’t like me.”</td>
</tr>
<tr>
<td>Stability</td>
<td>Stable: “I got a good grade because I’m good at that subject.” Unstable: “I got a good grade because I was lucky in the questions that came up.”</td>
<td>Stable: “I got a bad grade because I’m no good at that subject.” Unstable: “I got a bad grade because I hadn’t reviewed the material before the test.”</td>
</tr>
<tr>
<td>Specificity</td>
<td>Specific: “I’m good at that, but that’s the only thing I’m good at.” Global: “I’m good at that, so I’ll be good at everything.”</td>
<td>Specific: “I’m no good at that, but I’m good at everything else.” Global: “I’m useless at everything.”</td>
</tr>
</tbody>
</table>
he just had size. We didn’t have a lot of tall kids, and Leroy was 6–6, 6–7 . . . and Pop Herring thought we had plenty of guards but needed size” (Pickeral, 2009). Although technically Jordan was not “cut” from the varsity team, the effect on him was galvanizing, as he recalled in his acceptance speech when he was inducted into the National Basketball Association’s hall of fame on September 11, 2009: “Leroy Smith was the guy that when I got cut, he made the team [. . . ] he started the whole process with me, because when he made the team and I didn’t, I wanted to prove—not just to Leroy Smith, not just to myself—but to the coach who actually picked Leroy over me—I wanted to make sure you understood, ‘You made a mistake, dude’” (Jordan, 2009). In his junior and senior years, he did play for his high school’s varsity team, and played well, amassing over 1,400 points but not well enough to stand out nationally—at the end of his junior year, he was not on the list of the top 300 high school prospects being looked at by colleges (Bondy, 1999).

When Tom Brady was drafted into the National Football League in 2000, he wasn’t drafted until the sixth round (he was actually the 199th player selected) and was the 7th quarterback chosen that year.

Perhaps most remarkably, when Mike Piazza was drafted by the LA Dodgers in 1988, he was drafted 1,390th. In other words, the twenty-six Major League Baseball teams in existence at that time selected 1,389 players ahead of him (and only 43 players after him!). He was the 63rd player drafted by the Dodgers that year, and he wasn’t even drafted as a catcher (at the time, he was a first baseman). Realizing that he probably had a better chance of a major league career as a catcher, he began playing catcher in the Winter Leagues, and played twenty-one games for the Anaheim Angels in 1992 (in which he batted a modest .232). In the following year, however, he was the National League’s Rookie of the Year. Over the next fourteen years, he averaged over .300 and is now widely regarded as the best offensive catcher of all time, with career totals of 427 home runs and 1,335 runs batted in.

Each of these three individuals received feedback that they weren’t good enough, but each decided in the face of that feedback to improve rather than give up and do something else. The determination to do better was crucial in each of these cases.

Of course, whether a student sees feedback as relating to something that is permanent or transient depends on the student’s attitude. One
student may see feedback as sending the message “You're not smart enough,” while another student may see the same feedback as saying, “You're not smart enough—yet.” What matters is whether students see their future potential as limited by the current performance or not. Students who believe that ability is fixed will see any piece of work as a chance either to reaffirm their ability or to be shown up. If they are confident in their ability to achieve what is asked of them, then they will attempt the task. However, if their confidence in their ability to carry out their task is low, then they may well avoid the challenge (especially if they think others will succeed), and this can be seen in classrooms every day. A large number of students decide that, taking all things into account, they would rather be thought lazy than dumb and refuse to engage with the task. This is a direct consequence of the belief that ability is fixed. In contrast, those who see ability as incremental see all challenges as chances to learn—to get smarter—and, therefore, will try harder in the face of failure. These views of ability are generally not global—the same students often believe that ability in schoolwork is fixed and that ability in athletics is incremental, in that the more one trains at, say, triple jump, the more one's ability at that athletic endeavor increases. Therefore, what we need to do is ensure that the feedback we give students supports a view of ability as incremental rather than fixed: by working, you're getting smarter.

**A Recipe for Future Action**

All this suggests that providing effective feedback is very difficult. Get it wrong, and students give up, reject the feedback, or choose an easier goal. Even when the students engage with the feedback, there is still a danger that the focus is shifted to short-term rather than long-term goals. That is why I have become more and more interested in the practices of high-achieving sports coaches.

In many high schools, those involved in academic programs treat the school as a talent refinery. Their job is to deliver the curriculum to the students. Some students will get it and thrive, and others will not. Although most states require certain courses to be passed to qualify for a high school diploma, beyond that, if students cannot pass a course (such as Algebra II or trigonometry), then that's OK—it's obviously not their subject. In other words, the school functions rather like an oil refinery—its
job is to sort the students into different layers. Those involved in athletics programs cannot afford to do this. The football coach does not have the luxury of saying, “We don’t have a good enough quarterback, so we’re not going to play football this year.” Instead, the football coach sees his job as making the best quarterback they have as good as he can be, and doing the same at every other position. In other words, athletics coaches tend to see schools not as talent refineries but as talent incubators or talent factories. They see their job not as just identifying talent, but also nurturing it, and even producing it, often getting out of athletes more than the athletes themselves believed they could achieve.

Coaches do this through the provision of feedback that moves learning forward. In 1998, when Paul Black and I published “Inside the Black Box,” we recommended that feedback during learning be in the form of comments rather than grades, and many teachers took this to heart. Unfortunately, in many cases, the feedback was not particularly helpful. Typically, the feedback would focus on what was deficient about the work submitted, which the students were not able to resubmit, rather than on what to do to improve their future learning. In such situations, feedback is rather like the scene in the rearview mirror rather than through the windshield. Or as Douglas Reeves once memorably observed, it’s like the difference between having a medical and a postmortem.

Feedback functions formatively only if the information fed back to the learner is used by the learner in improving performance. If the information fed back to the learner is intended to be helpful but cannot be used by the learner in improving her performance, it is not formative. I remember talking to a middle school student who was looking at the feedback his teacher had given him on a science assignment. The teacher had written, “You need to be more systematic in planning your scientific inquiries.” I asked the student what that meant to him, and he said, “I don’t know. If I knew how to be more systematic, I would have been more systematic the first time.” This kind of feedback is accurate—it is describing what needs to happen—but it is not helpful because the learner does not know how to use the feedback to improve. It is rather like telling an unsuccessful comedian to be funnier—accurate, but not particularly helpful, advice.

Again, the differences in such practices between academic and
athletic programs are illuminating. Imagine a young fast-pitch softballer is struggling—her earned run average (ERA) is 10 (if you know nothing about fast-pitch softball, that’s not very good). If her coach were like many high school teachers, he might say, “Here’s my advice: you need to get your ERA down.” Accurate but not helpful. She knows she needs to get her ERA down, but she needs to know how.

The coach might look at her pitching and realize that the reason she is having a lot of runs hit off her is because she is trying, unsuccessfully, to pitch a rising fastball. This is a fastball thrown with so much rotation on the ball that as it reaches the plate, it rises sharply and is quite unhittable. Of course, if it doesn’t rise, it’s just a fastball over the middle of the plate, which is why she is getting hit for a lot of runs. So the coach says to the pitcher, “I know what’s going wrong. It’s your rising fastball. It’s not rising.” Again, accurate but not helpful.

But if the coach tells the pitcher that she is not dropping her pitching shoulder enough to deliver the pitch from below the knee, then this gives the athlete something to work with. The secret of effective feedback is that saying what’s wrong isn’t enough; to be effective, feedback must provide a recipe for future action.

This is actually inherent in the origin of the term feedback—a term borrowed from engineering. The term was first used by Norbert Wiener in 1946, and the important feature of feedback as it is used in engineering is that it forms part of a feedback loop. The classic example of a feedback loop is the regulation of the temperature in a room with the use of a thermostat.

The system has four key elements:

1. A means of setting the desired state (temperature setting)
2. A means of establishing the current state (thermometer)
3. A means of comparing the current state with the desired state (thermostat)
4. A means of bringing the current state in line with the desired state (furnace or cooling system)

For engineers, feedback about the discrepancy between the current state and the desired state is useless unless there is also a mechanism within the feedback loop to bring the current state closer to the desired state. In education, we use the term feedback for any information given
to students about their current achievements. Just telling the students that their current performance falls short of where they need to be isn’t feedback that an engineer would recognize. If there is nothing connecting the thermostat to the furnace, then there is no feedback loop and, therefore, to an engineer, no feedback.

To be effective as a recipe for future action, the future action must be designed so as to progress learning. In other words, the feedback must embody a model of progression, and here, again, is where much coaching in athletics programs is well designed. It is not enough to clarify the current state and the goal state. The coach has to design a series of activities that will move athletes from their current state to the goal state. Often coaches will take a complex activity, such as the double play in baseball, and break it down into a series of components, each of which needs to be practiced until fluency is reached, and then the components are assembled together. Not only does the coach have a clear notion of quality (the well-executed double play), he also understands the anatomy of quality; he is able to see the high-quality performance as being composed of a series of elements that can be broken down into a developmental sequence for the athlete. This skill of being able to break down a long learning journey—from where the student is right now to where she needs to be—into a series of small steps takes years for even the most capable coaches to develop. The best coaches say they are always learning how to be better at coaching. There are clear principles that can guide their development, which are discussed later in the chapter, together with some practical techniques for classroom implementation.

**Grading**

From the research discussed previously, it should be clear that the grading practices prevalent in most US middle schools and high schools are actually lowering student achievement. Worse, typical grading practices don’t even do the one thing they are supposed to do, which is tell us what students know:

Imagine, for a moment, a school that has an eight-week marking period, with students receiving a grade each week. Lesley starts out with four As but ends up with four Cs. Overall, of course, she gets a B. Chris, on the other hand, starts out with
four Cs but ends up with four As. He gets a B too.

But who has learned more? In terms of overall achievement, Chris, with his four final As, seems to have mastered the content for the marking period and really deserves an A. Conversely, Lesley, with her four final Cs, seems far from mastering the content, but she gets a B because of her good start. The fact is that our current grading practices don’t do the one thing they are meant to do, which is to provide an accurate indication of student achievement. (Clymer & Wiliam, 2006/2007, p. 36)

We can’t get rid of grades—the whole American system of college selection relies on them—but what we can do is design smarter grading systems that provide accurate information about student achievement while supporting student learning. The key to doing this is a principle outlined by Alfie Kohn (1994): “Never grade students while they are still learning” (p. 41). As soon as students get a grade, the learning stops. We may not like it, but the research reviewed here shows that this is a relatively stable feature of how human minds work. So we have to deal with it and design assessment systems accordingly.

If grades stop learning, students should be given them as infrequently as possible. In high school, there may be an argument for one per marking period but certainly no more. In middle school, there may be a case for grades once a year, but in elementary school, the use of grades appears to be entirely unjustified.

Many administrators realize this but continue to mandate grades because they believe that parents want them, and surveys of parents often show support for grades, but this is hardly an informed choice. Parents often support grades only because they want to know how their children are doing in school and assume that grades are the only way to find out. However, parents have no idea what their children’s grades mean, because, as Paul Dressel remarked over half a century ago, “A grade can be regarded only as an inadequate report of an inaccurate judgment by a biased and variable judge of the extent to which a student has attained an undefined level of mastery of an unknown proportion of an indefinite material” (Dressel, 1957, p. 6). Once we have them hooked, students want grades, too, but rather than allowing them to become codependents in an unhealthy relationship, it would surely be
better not to get students addicted to grades in the first place.

We need classroom assessment systems that are designed primarily to support learning and deal in data that are recorded at a level that is useful for teachers, students, and parents in determining where students are in their learning. Such fine-scale evidence can always be aggregated for summative reporting. It is not possible to go the other way: from aggregate reports of achievement to learning needs.

For example, a swimming teacher was observing students in a swimming pool. She had drawn a grid listing the students’ names in the first column. The other four columns were headed: arms, legs, breathing, and timing. As she observed each student, she entered a 0, 1, or 2 in each column, according to the level of competence displayed by the student in that aspect of swimming. She could, if she wanted to, add up the scores and give each student a total out of eight, but knowing that someone has scored seven out of eight is useless for planning instruction. On the other hand, knowing that a student’s arms, legs, and breathing are fine but the timing requires attention is useful to the teacher (and would also be useful to another teacher if one had to take over teaching this group).

Middle school science specialist Jacki Clymer and I described how these principles could be implemented in a science classroom (Clymer & Wiliam, 2006/2007). For each marking period, the key learning outcomes are identified. For example, for the first marking period, a total of ten areas of interest are adopted (some of these might cover multiple learning outcomes):

1. The appropriate use of laboratory equipment
2. Metric unit conversion and labeling
3. Calculating density
4. Applying density (floating, sinking, layering, thermal expansion)
5. Density as a characteristic property
6. The phases of matter (at a molecular level)
7. Gas laws
8. Communication (graphing)
9. Communication (lab reports)
10. Inquiry skills
For each of the ten areas of interest, sources of evidence are identified. For example, the evidence for the appropriate use of laboratory equipment comes from observations, from a homework assignment on safety, and from the laboratory reports completed by the students. On the basis of the evidence collected, the teacher scores each area for each student as either 0 (no evidence of mastery), 1 (some evidence of mastery), or 2 (strong evidence of mastery), and keeps these data in a spreadsheet. The teacher uses the “conditional formatting” feature in the spreadsheet to highlight a cell in green if there is a 2 in the cell, yellow if there is a 1, and red if there is a 0. The result is an instant summary of the achievement of the class, as shown in figure 5.2 (page 126).

Since there are ten areas for this marking period, each student’s total is divided by twenty to yield a percentage, which can then be used to provide an indication of what the grade would be if the marking period ended at that point. However, at any time before the end of the marking period, students can provide further evidence of their competence. So, for example, if Emma Bettany wants to know what she needs to do to get an A, the teacher can quickly tell her that all she needs to do is provide evidence that she has mastered one of the three areas for which the teacher has, to date, seen only evidence of partial mastery.

At the end of the marking period, the students take a test, which is used to confirm the evidence collected up to that point. When the performance on the test indicates a different level of mastery from that indicated on the basis of the class work, this is treated as a matter needing further investigation (rather than, for example, just averaging the scores). Typically, the teacher will interview the student to probe the extent of the student’s understanding of the matter in question.

Perhaps the most profound impact of such a grading system is that it pushes both teacher and students into thinking about longer-term learning. If a student shows mastery of something at the beginning of the marking period but then fails to do so later, his grade can go down. This is exactly what we found in a study of the implementation of this system with an eighth-grade class. Students became more engaged in monitoring their own learning; frequently asked for clarification, both from the teacher and from their peers; and regarded the teacher more as a coach than a judge. Their achievement went up, too (Clymer & Wiliam, 2006/2007).
<table>
<thead>
<tr>
<th>Period</th>
<th>Use equipment</th>
<th>Metric units</th>
<th>Calc. density</th>
<th>Apply density</th>
<th>Density property</th>
<th>Phases of matter</th>
<th>Gas laws</th>
<th>Comm. (graphs)</th>
<th>Comm. (reports)</th>
<th>Inquiry skills</th>
<th>% Grade</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allen</td>
<td>James</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>75%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Ayers</td>
<td>Liam</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>65%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Baldwin</td>
<td>Lee</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>75%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Bettany</td>
<td>Emma</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>85%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Birch</td>
<td>Leah</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>65%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Burns</td>
<td>Robert</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>50%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Cobern</td>
<td>David</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>80%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Creasey</td>
<td>Simon</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>65%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Darby</td>
<td>Hannah</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>80%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Eastwood</td>
<td>Luke</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>60%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Ferguson</td>
<td>Mark</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>85%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Forbes</td>
<td>Sarah</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>40%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Goodger</td>
<td>Mark</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>95%</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Hall</td>
<td>Mark</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>85%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Howells</td>
<td>George</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>70%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Hudson</td>
<td>Kisty</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>Hurley</td>
<td>Victoria</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>50%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Langan</td>
<td>Jennifer</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>90%</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Larkin</td>
<td>Andrew</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>55%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Leach</td>
<td>Jonathan</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>75%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Lowings</td>
<td>Charlotte</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>75%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Mcglashan</td>
<td>Scott</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>50%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Parr</td>
<td>Amy</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>75%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Ringham</td>
<td>Grace</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>40%</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Rosamond</td>
<td>Lee</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>70%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Rose</td>
<td>Peter</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>80%</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Ryder</td>
<td>Thomas</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>70%</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Sweats</td>
<td>William</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>90%</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Walton</td>
<td>Emma</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>100%</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Average facility</td>
<td>76%</td>
<td>86%</td>
<td>71%</td>
<td>84%</td>
<td>98%</td>
<td>78%</td>
<td>88%</td>
<td>69%</td>
<td>29%</td>
<td>31%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.2: Grading scheme based on Clymer and Wiliam, 2006/2007.

© 2011 Solution Tree Press. All rights reserved.
A particularly important feature of this grading system is that it avoids the ratchet effect present in most grading systems—the idea that a student’s grade for a particular piece of work can never go down. When students are allowed multiple resubmissions and can improve work without penalty, this can have negative consequences. For example, in many high schools, students know they can turn in very poor quality work, get feedback on it, and improve it for a passing grade. The problem with such systems is that there is no incentive for students to make the first submission as good as it can be, because it can always be remediated in a second submission. Districts that allow students to resubmit work multiple times need to consider creating systems that avoid such problems. Deanna Holen, a physical education teacher in Chico, California, designed a grading rubric that was constructed in such a way that no matter how poor the original submission, a resubmission could bring it up to a passing grade, but it also ensured that the better the first submission was, the higher the final grade would be, as shown in table 5.4 (page 128).

Another way to provide similar incentives is to allocate 50 percent of the available points to the first submission and 50 percent to the improvement shown in the work as a result of responding to the feedback.

Some teachers question whether we should ever be satisfied with any work that is not of a high standard. Joe Rubin, a science teacher in San Francisco, only gives one grade: an A. When he grades work, it either comes back with an A or is returned with a note saying it is not yet ready to be evaluated. Of course, the students who need multiple attempts to get an A may well run out of time, and thus complete fewer assignments, but the message this teacher is sending to all students is that only the best is good enough. More importantly, he sends the message that everyone can do quality work, although some need more support and guidance than others. He emphasizes quality over quantity.

**Practical Techniques**

If I had to reduce all of the research on feedback into one simple overarching idea, at least for academic subjects in school, it would be this: feedback should cause thinking. All the practical techniques discussed here work because, in one way or another, they get the students to think, rather than react emotionally to the feedback they are given. When I talk to teachers about the first of the two studies by Ruth Butler...
and her colleagues discussed previously (the one in which students were given comments, scores, or both), I ask them, “When the students were given both comments and scores, which did they look at first?” Everyone realizes, of course, that it was the score. What is more interesting is what the students looked at next: somebody else’s score.

As soon as students compare themselves with someone else, their mental energy becomes focused on protecting their own sense of well-being rather than learning anything new. In one school I have worked with, the principal insists that there is a score or a grade on each student’s work every two weeks. The language arts teachers give students the traditional letter grades at the end of each marking period, but during the marking period, they use a grading scheme that they call “minus, equals, plus.” As well as comments, each student receives a symbol of –, =, or + depending on whether the work submitted was not as good as, about the same as, or better than his or her last work on the topic. The students who used to get A grades hate it, because they have to keep getting better to get a +. It is particularly interesting when high-achieving students in these classes compare their grades with other students’ and find that they have received = or even – when students with lower absolute achievement get a +. The important feature of such a grading scheme is that it feeds back to learners about things that are within their control, such as whether they are improving, rather than things that are not within their control, such as how they compare with other students in the class.

To be effective, feedback needs to direct attention to what’s next rather than focusing on how well or badly the student did on the work, and this rarely happens in the typical classroom. Pam Hayes, a middle

Table 5.4: Grading Scheme to Encourage Both Improvement and Good First Attempts

<table>
<thead>
<tr>
<th>Grade on first submission</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>
school math teacher, told me about a conversation she had with a fifth grader in which the student said, “When you get a lot of feedback on your work, it means it wasn’t very good.” When asked to explain this rather surprising comment, the girl pointed out that successful work is usually just given a (high) grade and a comment like “Good job,” whereas less successful work is returned to the student with lots of annotation from the teacher. To this girl, the more “feedback” you got, the worse your work must have been. In many classrooms, teachers require students to do corrections for homework, leaving high achievers with nothing to do. Used in this way, feedback really is punishment.

If, however, we embrace the idea of feedback as a recipe for future action, then it is easy to see how to make feedback work constructively: don’t provide students with feedback unless you allow time, in class, to work on using the feedback to improve their work. Then feedback is not an evaluation of how well or how badly one’s work was done but a matter of “what’s next?”

One technique for structuring this that is particularly effective when responding to a piece of student writing is called “three questions.” As the teacher reads each student’s work, when she sees something on which she would like the student to reflect, she places a numbered circle at that point in the text. Underneath the student’s work, the teacher writes a question relating to the first numbered circle, leaves a number of lines for the student’s response, writes a question for the second, leaves space for the student’s response, and then writes a third question. The first ten or fifteen minutes of the next lesson are taken up with responding to the three questions posed by the teacher. The important feature of this technique is that no matter how bad or good your work was, everyone has the same amount of work to do.

This idea that feedback is about “what’s next?” also addresses another shortcoming of much current practice. I often ask teachers whether they believe that their students spend as much time utilizing the feedback they are given as it has taken the teacher to provide it. Typically, fewer than 1 percent of teachers believe this to be the case, and this needs to change. The first fundamental principle of effective classroom feedback is that feedback should be more work for the recipient than the donor.

Giving students comments, rather than grades or scores, is obviously useful, but most teachers still find it difficult to get the students to read
the comments. Charlotte Kerrigan, a language arts teacher, had been giving only comments to her students for a while but was still unhappy with the amount of attention the students were giving to her comments, so she made a small but extremely powerful change in the way she provided feedback. Her tenth-grade class had just completed essays on a Shakespeare play they had been studying. Kerrigan collected the essays, and instead of writing her comments in the students’ notebooks, she did so on strips of paper. Each group of four students received their four essays and the four strips of paper, and the group had to decide which comment belonged with which essay.

A second principle of effective feedback is that it should be focused. We generally give our students large amounts of what we call feedback (although an engineer probably wouldn’t), but it is usually of moderate quality, and we generally don’t require students to do much with it. In giving feedback, less is often more.

I learned this the hard way when I first became a teacher educator, working with preservice teachers and observing them during their periods of practical teaching. I would sit at the back of their classrooms and assiduously note all the errors they had made in their teaching—I would frequently generate as many as three pages of comments on all the errors the student-teacher had made in a forty-five-minute lesson. At the end of the period, I would bestow this wonderful feedback on the hapless student. I was frustrated that my wonderful feedback did not seem to be having any effect. After a while, I realized that the problem was that I was giving far too much feedback; I needed to give less, but more focused, feedback. Rather than handing over a whole list of errors to be corrected, I started saying things like, “Over the next two weeks, I want you to work on these two things: first, before you give any important instructions to the class, make them put their pens down, and second, make sure that you don’t talk over the students—if they start talking to each other when you are talking, stop talking.” I gave less feedback and had more impact by being more focused.

A third principle is that the feedback should relate to the learning goals that have been shared with the students. If the teacher has provided a scoring rubric, then it is important that the feedback relates to that rubric. If there are learning intentions and success criteria for the work, then the feedback should loop back to those. This sounds obvious, but I have lost count.
of the number of times I have seen teachers provide students with rubrics or success criteria and yet fail to use these in framing their feedback to the students.

Math teachers may be thinking that ideas such as comment-only grading may work well in language arts and social studies, but math is different. After all, if a student has solved twenty equations, and the teacher places a check mark next to fifteen of them and a cross next to the other five, the student can work out his score of fifteen out of twenty, even if the teacher does not put a score of 75 percent on the work. As noted previously, however, what is important is not the form that the feedback takes but the effect it has on students.

Putting a check or a cross next to each of the solutions leaves nothing for the student to do, except maybe correct those that are incorrect. An alternative would be to say to the student, “Five of these are wrong. You find them; you fix them.”

This technique is particularly well suited to mathematics as it is often easier to check whether a solution is correct (for example, by substituting for a solved variable back into the original equation). However, it can also be readily adapted to other subjects. When reviewing a final draft of a piece of writing, a language arts teacher placed a dot in the margin of each line where there was something that needed attention. For weaker students, she replaced the dot with a g for an error in grammar, an s for a spelling error, a p for punctuation, and so on, thus differentiating the feedback. The important point is that the feedback is focused, is more work for the recipient than the donor, and causes thinking rather than an emotional reaction.

**Conclusion**

The word *feedback* was first used in engineering to describe a situation in which information about the current state of a system was used to change the future state of the system, but this has been forgotten, and any information about how students performed in the past is routinely regarded as useful. It is not. In this chapter, we have seen that in almost two out of every five carefully designed scientific studies, information given to people about their performance lowered their subsequent performance. We have also seen that when we give students feedback, there are eight things that can happen, and six of them are bad (table 5.2, page 115).
Some ways to give effective feedback have been described in this chapter, but every teacher will be able to come up with many more, provided that the key lessons from the research on feedback are heeded. If we are to harness the power of feedback to increase student learning, then we need to ensure that feedback causes a cognitive rather than an emotional reaction—in other words, feedback should cause thinking. It should be focused; it should relate to the learning goals that have been shared with the students; and it should be more work for the recipient than the donor. Indeed, the whole purpose of feedback should be to increase the extent to which students are owners of their own learning, which is the focus of the next two chapters.
EmbEddEd formative assessment

Dylan Wiliam's wonderful new book strips away the murk currently surrounding formative assessment. Deftly moving between research evidence and classroom practicalities, Wiliam provides a compelling vision of how formative assessment can enhance students' learning. For today's educators, this is truly a must-read! —W. James Popham, Professor Emeritus, UCLA

"With precision, clarity, and practical comprehensiveness, Wiliam gives us all we need to know to get deep and lasting success. Each of his five strategies is a gem; in combination, they are unbeatable.” —Michael Fullan, Professor Emeritus, University of Toronto

If we want our students to thrive in the complex, unpredictable world of the 21st century, we must concentrate on increasing educational achievement by increasing the quality of the teachers in our schools. *Embedded Formative Assessment* faces this challenge head-on by making a case for the important role of formative assessment as a process, not a tool, in increasing teacher quality and thus student learning.

Author Dylan Wiliam:

- Provides over fifty practical techniques for classroom formative assessment
- Emphasizes the instructional side of the process
- Explores in-depth the use of classroom questioning, learning intentions and success criteria, feedback, collaborative and cooperative learning, and self-regulated learning to engineer effective learning environments

"This is a must-have, must-read, must-put-into-immediate-action book for every teacher, principal, and network leader. Dylan has made powerful coaching learning ideas accessible in a highly engaging way and makes the development of a learning system possible.” —Linda Kaser, Co-Director, Network of Performance Based Schools, and Co-Director, Center for Innovative Educational Leadership, VIU

solution-tree.com