



Rodney Y. Stutzman and Kimberly H. Race

Assessment

EMRF: Everyday Rubric Grading

TEACHERS SPEND CONSIDERABLE TIME AND energy working to clarify and communicate their expectations to students. Those expectations are typically communicated when teachers present assignments. Reinforcing the assignments with an assessment process that narrows the feedback focus can benefit teachers and enrich students. For individual assignments, students really need to know the answers to the following questions: Is my work at an acceptable level? Am I on track to earn the grade to which I aspire in this course? They do not need feedback carved into razor-thin slices to make those determinations. We looked for an alternative that allows us to effectively evaluate responses to rich problems within a standards-based curriculum. We hoped to be able to assess written communication, as well as mathematical computation. Promoting student responsibility with a student-centered system was also important. Ultimately, a revised assessment system, which we have come to call *EMRF*, emerged.

The NCTM's *Assessment Standards for School Mathematics* (NCTM 1995) suggests the following considerations in an assessment system:

- Assessment should reflect the mathematics that all students need to know and be able to do.
- Assessment should enhance mathematics learning.
- Assessment should promote equity.
- Assessment should be an open process.
- Assessment should promote valid inferences about mathematics learning.
- Assessment should be a coherent process.

To incorporate these considerations, we have found that using a four-tier rubric to evaluate students' work is helpful. We suggest that establishing the four levels is beneficial to both the teacher and the student. Students ultimately want to know the following: Is this an example of excellent work? Am I meeting the expectations for the material? Do I need to go back and revise or relearn some of this material? Do I even have pieces of understanding here?

A few good-quality questions can allow teachers to determine and report back who has partial understanding of the mathematics, who has a solid understanding, and who understands in depth. Teachers do not need to worry that they really need a certain number of questions to generate enough data to give a valid grade. A well-written prompt and rubric set creates opportunities for assessments with questions that are fewer in quantity but higher in quality. The teacher can provide helpful feedback to students with a smaller number of questions.

THE FOUR LEVELS

Our original attempts to communicate feedback on a four-tier rubric were confounded by the very concept of percents. We wrote descriptors of four levels, with level 4 demonstrating excellence and level 1 signifying a fragmentary attempt. The levels were not intended to represent a linear scale. Each step signified a considerable improvement in understanding, yet students still saw a 3 as meaning "three out of four," or 75 percent. They would take a *revisions needed*—2 grade, average it with an *excellence*—4 grade, and see the grade as equivalent to getting two *meets expectations*—3's even though they passed one assignment with flying colors and almost failed the other.

To prevent this averaging, we shifted to the letter codes E, M, R, and F. An E signifies *excellent*, M is *meets expectations*, R means *revisions required*, and F is *fragmentary*. The gap in understanding demonstrated by an M over an R is greater than

Rodney Stutzman, stu_tz_mans@hotmail.com, teaches mathematics at Littleton High School, Littleton, CO 80121. He is currently on an extended leave while living in Tanzania and working with Mennonite Central Committee, a relief and development organization. Kimberly Race, krace@lps.k12.co.us, also teaches mathematics at Littleton High School. She is interested in alternative forms of assessment and has done extensive professional development in standards-based curricula.

Students do not need feedback carved into razor-thin slices

the refinement and generalization demonstrated by an E over an M. The black bar in the flowchart in **figure 1** signifies the important division between demonstrating understanding and not demonstrating understanding. Using letters rather than numbers to represent levels of understanding helped students move away from the linear percent scales to which they had grown accustomed.

WRITING RUBRICS: DEFINING THE MATHEMATICS THAT ALL STUDENTS NEED TO KNOW

We learned to begin writing specific assignment rubrics at the M (meets expectations) level. Major considerations are, What do I expect a student to demonstrate with this prompt? and What evidence does the student need to show to convince the reader that he or she understands? The rubric at the M level is typically a checklist of basic expectations. Those expectations reflect both the mathematical content and communication. (Communication is the students' description of the process that they used to solve a problem and justify their answers.) An M does not represent perfection. It is an attempt to describe indicators of a solid, but basic, understanding of the material. Procedural understanding is evident at this level. Conceptual understanding is probably formative but not mastered. Likewise, communication is adequate but not refined. We often refer to level M as *first-draft quality*. Any errors are minor and could be corrected by the student in another draft. Typically, we do not request that draft—the student has already demonstrated understanding. The student needs no assistance beyond written feedback pointing out any errors.

After addressing the expectations for an M, we address the expectations for an E (excellence). Which indicators demonstrate that the student has risen above a basic, solid understanding to a level of excellence? What defines excellence in the communication of students' thoughts?

Typically, the first statement in the E description is, "meets all expectations of an M." Level-E work shows generalization and clear communication. It is final-draft quality on a first attempt. The work is accurate—the reader is not distracted by errors. It may contain trivial errors; but as a whole, the work is representative of excellence. The student has demonstrated solid conceptual understanding and can defend his or her result and any generalizations with a logical, written explanation.

After preparing descriptors of M's and E's, we move to describing level R (revision required). Instead of a checklist, these descriptors tend to represent a series of tripwires: What common misunderstandings you are concerned about? What errors show that a student "does not have it"? What are typical errors that students need to overcome to be

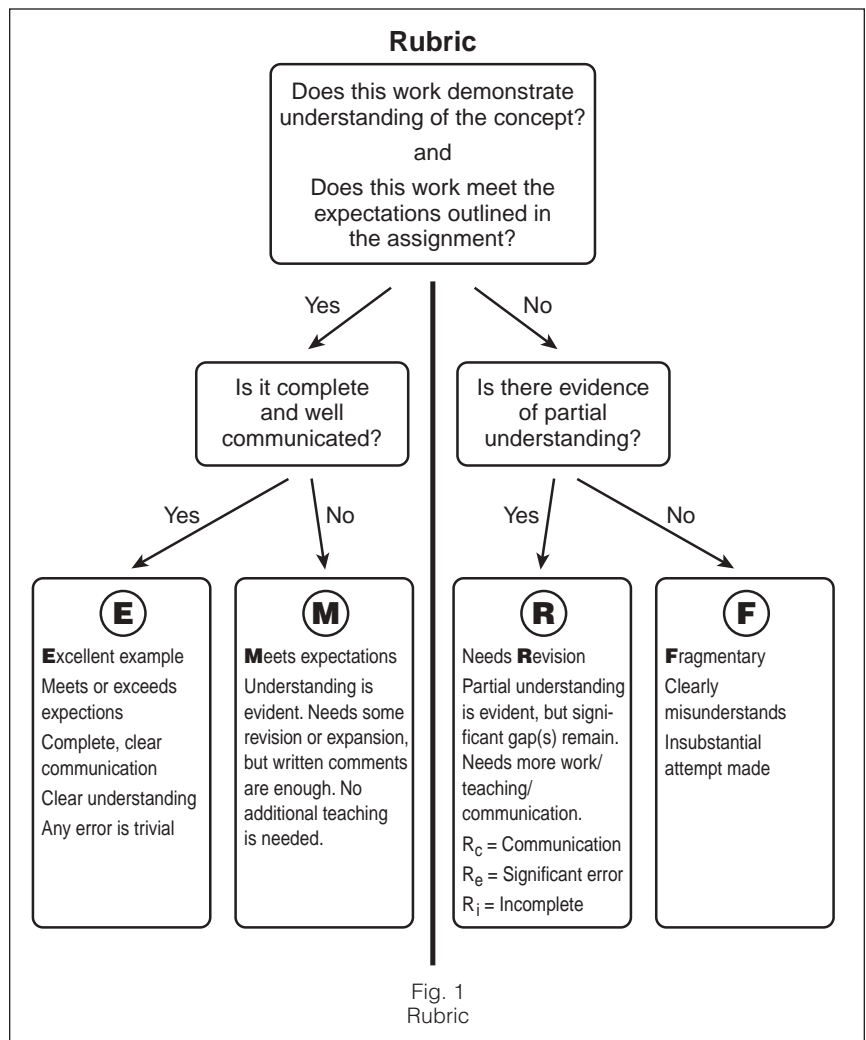
successful? Normally, the errors evident in level-R work require some additional teaching, either from the teacher or from a peer. The student is likely to need more assistance than written comments alone can provide.

Communication concerns can also trigger a categorization of R. The mathematics might be correct; but the communication is nonexistent, garbled, or trivial. The write-up demands strong background knowledge on the part of the reader, since an outsider would not be able to audit the work or follow the thought process.

The descriptor of an R might also address "not all questions attempted" or "lack of investigative data." The piece might demonstrate a substantial level of understanding but be incomplete. Students have a responsibility to complete the prompt in its entirety.

Level R tends to be the broadest of the levels. It covers work that does not yet clearly demonstrate understanding but does show evidence of at least partial understanding. Given the range of responses

Writing a rubric defines a standard



It became common to hear students say, “No, this system makes sense”

that can trigger an R, we have added subscripts to help direct the students to the area of weakness or concern. R_c signifies that the *communication* is insufficient. R_i tells the student that the response is *incomplete*, whereas R_e alerts the student to a computational or conceptual *error*. These subscripts are not intended to replace teacher comments; they are intended to help clarify and remind students of teacher expectations for communication, completeness, and accuracy.

The last level is F (fragmentary): the student clearly made some attempt; but the work is fragmented, completely misdirected, or unsubstantiated. It shows no clear evidence of even partial understanding. More precisely, the work lacks the basic foundation, structure, or framework for revision. The rubric rarely says more than “fragmentary,” since predicting such profound errors is difficult.

Figure 2 gives a sample prompt and its accompanying EMRF rubric.

Writing a rubric defines a standard. The very process asks the teacher to clarify his or her expectations before grading papers. Therefore, the teacher should make every effort to have a complete and well-thought-out rubric before he or she reviews students’ responses. Such a rubric helps the teacher avoid the “urge to curve.” A preponderance of R’s is valuable information for planning upcoming lessons. If students do not meet the standard, artificially adjusting the grading scale does not change their level of understanding. Devoting some class time to letting students work through, or at least begin, their revisions in class might be more effective. Students then have the power to adjust their grade upward while significantly increasing their level of understanding. Teachers are put in a position where they can draw valid inferences from students’ work. The feedback record represents student understanding as measured against a carefully defined standard.

SHARING THE RUBRIC PUBLICLY: AN OPEN PROCESS

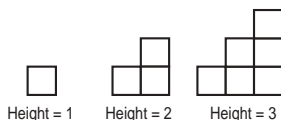
We have found that sharing the specific rubric for an assignment when returning papers is important and helpful because doing so clarifies and refines the teacher’s expectations for students. They need not speculate about the quantity of partial credit. It creates a teachable moment within the mundane administrative task of returning papers. Students are informed by seeing clearly described expectations and standards. Sharing the rubric promotes student reflection and self-assessment. When the rubric is posted and students have their papers in hand, they cannot help but reflect on the question, Did I meet the standard? Posting the rubric also promotes a sense of fairness and openness in grading. It clearly indicates that each student is working toward the same standard; when the rubric is open and available for all to consider, there is no sense of favoritism or shifting standards.

Sometimes, sharing the rubric when the assignment is given promotes quality work. Students know exactly what is expected. They are not put in the position of miscalculating the answer to the question, What is the minimum that I can get away with? They are told in advance what they need to do to meet expectations and what they need to do to demonstrate excellence. When the teacher shares the rubric in advance, avoiding “giving away the answer” can be difficult and might require a deliberate vagueness. For example, the descriptor might be “has correct answer” rather than “1808 rats.”

REVISIONS: PLACING RESPONSIBILITY AND CONTROL WITH STUDENTS

The rubric indicates work that does not demonstrate

Consider the following pattern of stair steps:



If the pattern continues—

- How many blocks will be required for a height of 93?
- How many blocks will be required for a height of 748?
- How many blocks will be required for a height of n ?
- A staircase used 2556 blocks. How tall was it?

(a)
Sample prompt

E:

- All parts of an M are included.
- Author was able to answer part (d) correctly ($n = 72$).
- Part (d) is well communicated.

M:

- There is a rule for question (c).
- That rule allowed the author to answer questions (a) and (b) correctly:
(a) 4371
(b) 280,126
- The author attempted (d) but did not solve it.
- The write-up not only has the solutions—it also explains the thinking or process clearly. Any reader can understand what the author did and why she or he did it.

R:

- The author provided solutions; but no write-up, description, or explanation was present.
- The write-up is sketchy; the reader has to make assumptions about what the author was thinking.
- Not all questions were attempted.
- Solutions are given for (a) and (b), but no rule is provided for (c).
- There is a rule for (c), but it is not supported or explained.

F:

- Fragmented response, minimal attempt, no process evident, leaves reader with far more questions than answers.

(b)
Sample rubric

Fig. 2

understanding with the grade “R—revisions needed.” We have found that allowing students to actually revise that work is important. Many management systems are possible, but we maintain four rules for the revision process.

First, the revised work must be on a separate sheet of paper, so that any reader can see what was original and what constitutes a revision.

Second, the student must explain his or her original error. The explanation should tell the reader something that is not already obvious. Saying “I got the wrong answer” is not enough. It is more helpful to report, “When substituting into $a^2 + b^2 = c^2$, I put the hypotenuse into a leg spot,” or “I didn’t realize that there was a negative solution.” The error explanation should demonstrate that the student has investigated the original errors and is able to address them in some detail. The error explanation makes each student’s work unique. The intent is to prevent a student from simply seeking accurate work to copy. By insisting that students describe and debug their original errors, their revisions are specific to their situation. They cannot address their original attempt without having an improved understanding. The explanation of the error should be detailed enough to convince the reader that the student is unlikely to repeat that error.

Third, the revision should include a correct working of the problem or an explanation of it. That rework should be a “complete, perfect, excellent, top-to-bottom, left-to-right, start-to-finish” version of the problem, even if the error occurred late in the original attempt.

Fourth, we use a revision due date of two days from the time that papers are handed back. We thereby avoid the problem of having panic-stricken students who want to turn in several revised papers in the days immediately before grades are due. The teacher’s time dedicated to checking revisions is balanced throughout the grading period. Since students complete their revisions in a timely manner, a student with a level-R understanding can continue to move through the curriculum. As the students move on, they also have the opportunity to see new connections with the concepts with which they struggled.

Teacher discretion should be used when deciding how many times a student can work on a revision. The intent of allowing revisions is to enable the student to investigate the error with the help of peers or the teacher. Most students turn in successful revisions on the first attempt. We find that checking revised work does not take long. The teacher has recently read the original responses. The revised responses, which are generally complete and correct, take far less time to check than the original attempts that required conventional comments.

The essential task of the student is to demonstrate, “Yes, I made a mistake originally, but this work proves that I now understand.” We often describe it as analogous to winning a case on appeal: a person may lose the first trial, but if he or she is willing to do some paperwork and improve his or her arguments, he or she has the chance to carry the day in appellate court. The error explanation and detailed reworking of the problem constitute the student’s appeal. It is the student’s chance to demonstrate understanding. A student who is not able to demonstrate understanding until Tuesday when most of the class met that requirement the previous Friday is not denied credit for his or her achievement.

Revisions are another open process. Students must be provided with extra assistance; and they are encouraged to seek help from peers, tutors, teachers, parents, and so on. The R level indicates that students did not demonstrate understanding and that additional teaching is necessary. However, that teaching does not need to be provided in a whole-class setting—or even by the teacher—but students should have access to additional help. Revisions make assessment part of a coherent learning process. Formal assessments are no longer the end of the process for students who struggle.

Figure 3 is a sample of level-R work on the stair-step problem. **Figure 4** is the same student’s revised work at level M.

Students who successfully revise their work can improve their grade. If we are going to hold high standards for levels E and M and if we wish to avoid the temptation to curve, the students need an alternative route to enable them to have control over their semester grades. A revision process puts pressure on students to follow through with revisions instead of placing pressure on the teacher to reduce his or her standards and expectations with curved grades.

We have decided that work originally graded at level R can be revised to level M with high-quality revisions. Students often ask why an R cannot be upgraded to an E. After deliberation, we realized that part of our definition of excellence is the demand-task aspect. In essence, we are asking students to sit down and demonstrate their understanding within thirty minutes. Students who produce work at the excellent level are meeting that demand. That aspect cannot be reproduced within the revision process. Teachers might choose to allow students to obtain E’s through the revision process for prompts that are not demand-task appropriate.

Permitting revisions on work that originally received a mark of F is evaluated on a case-by-case basis. A brief interview with the student can often give the teacher insights on the cause of an F. The intent is to avoid creating a system in which the demand-task characteristic of formal assessment is

*Revisions
make
assessment
part of a
coherent
learning
process*

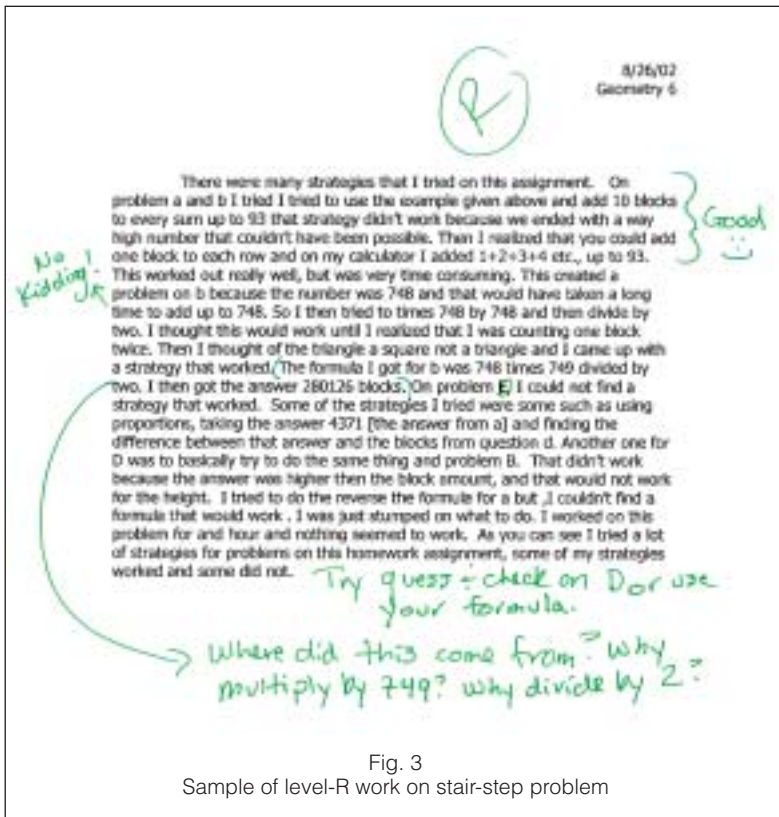


Fig. 3
Sample of level-R work on stair-step problem

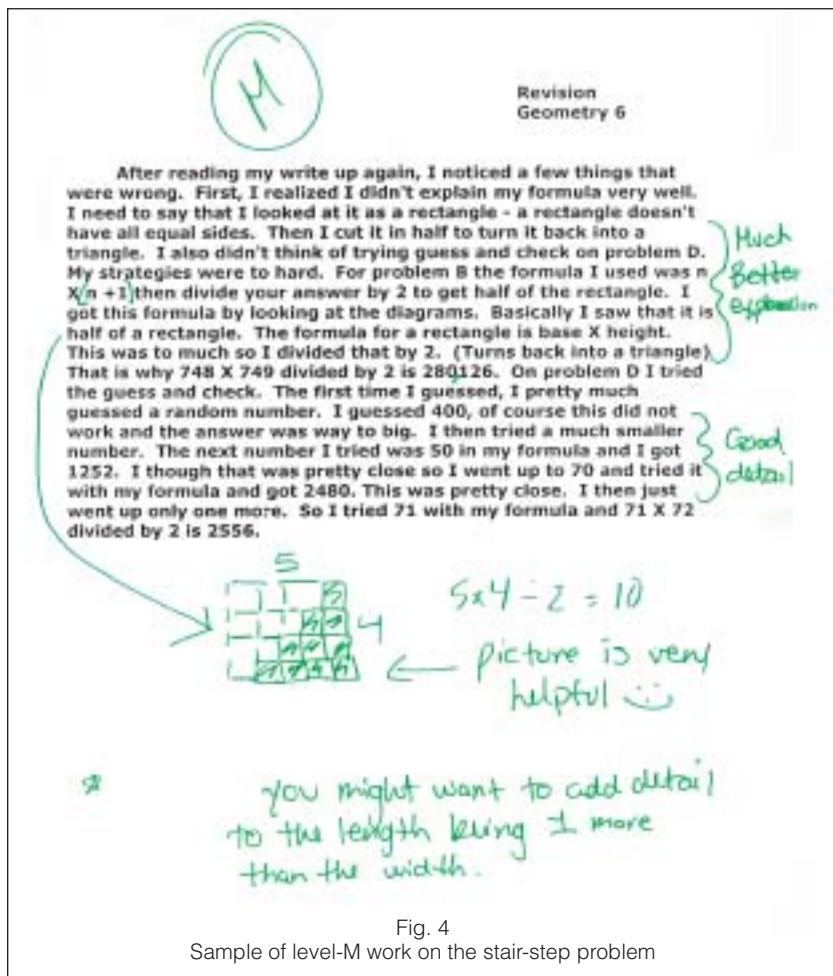


Fig. 4
Sample of level-M work on the stair-step problem

circumvented by the revision process. The revision process is intended to give students the opportunity to refine work in progress. It is not intended to move assessment entirely out of the classroom setting.

We typically expect the bulk of the revision write-up to be completed outside class. We have found that allowing some class time to begin the process when papers are returned is helpful. Students who need to revise work typically appreciate the opportunity to investigate their errors immediately. Those who were successful with the prompt are in demand, and are engaged, as peer coaches. The class time devoted to this investigation stage is highly productive, but the majority of the actual revision write-up is done outside class.

Interestingly, we have found that the revision process has several unexpected benefits. Often, time is an issue for special-needs students. The revision process gives them an opportunity to work on pieces in class under normal time constraints and to have an opportunity to demonstrate their understanding in an untimed setting. When parents discover that students have been given control over their grades through the revision process, they find it a pleasant surprise. Ultimately, students who are willing to invest in the process do not have failing marks in their grade record. The burden of achieving low grades is transferred to the student. In addition, high-achieving students are pleased that their goals and expectations are not put at risk by one poor grade on an examination.

SEMESTER GRADES

Teachers are ultimately asked to condense a semester's worth of work into a single letter grade. Clarifying and communicating expectations for semester grades requires a process that is similar to clarifying and communicating expectations for individual assignments. Teachers need to describe in advance what they expect of an A student, what they expect of a B student, and so on. At this point, we do not entirely avoid the use of ratios or percents to help in those semester-grade descriptions. Ratios are comfortable for parents and administrators. Ratios are more appropriate in evaluating the large amount of data represented by a student's semester record than they are in evaluating a single response to a prompt.

Semester grades are not calculated in a vacuum—student presentations, daily homework, group interactions, and other factors are typically considered in a grading plan. Teachers generate expectations on the basis of their experience and circumstances.

In evaluating the written-work component, we look at the number of concepts passed during the semester. Both E's and M's indicate that the student understood, or passed, the concept. We make

no distinction between M's received on original attempts and those received through the revision process.

A semester grade of A might require unit assessments that are consistently at level M or E, other assessments (quizzes, assessed homework) consistently at level M or E (most at level E), and no missing pieces. The description of a B might indicate that at least 75 percent of all assessments are level M or E and that at least one-fourth of those pieces are E's. Requirements for a C might be that a clear majority (65 percent) of all assessments are at level M or E.

In short, A students work to be sure that they receive a mixture of E's and M's, with at least half being E's. B students pass at least 75 percent of the concepts, and C students pass 65 percent of the concepts.

Grade computation becomes a matter of counting E's and M's rather than computing point totals. For example, if students have been presented with twenty concepts in assessments, a student who receives a B would have passed at least fifteen of them. Since students also have a chance to revise work with which they originally struggled, it might not be unreasonable to require A students to obtain eighteen or nineteen E's and M's, either as original marks or through the revision process. We have found that a grade line of E's, M's, R's, and F's is no more difficult to handle than a line of points. Grading software can be adapted to assist with the counting.

CHALLENGES

Students initially resisted the move from assessment on a linear scale to assessment against a defined standard. They experienced the discomfort associated with change. They pressed for a return to the familiarity of a continuous scale. Time was needed to communicate the meaning and intent of the EMRF system. However, when students became accustomed to the system and saw the benefits available to them, the tone of the comments changed. Hearing students say, "This system makes sense" became common.

The discipline to write rubrics for each piece was another initial challenge. With practice, it became apparent that the habit was actually saving time. Rather than unconsciously reconsidering expectations in moving from paper to paper, we found ourselves referring to the written rubric and evaluating papers more quickly and consistently. As previously mentioned, having that rubric available to then share with students had many benefits.

CONCLUSION

The EMRF system supports problem-based, standards-based curricula. It helps define expecta-

tions for student work and helps teachers create an open, equitable, student-centered process. The four-tier rubric is designed to convey essential information to the student. The four levels are carefully defined, and their descriptors are available to the students. Assessing work against a standard and incorporating a revision process help teachers collect information and make inferences to assist lesson planning, while placing considerable responsibility and semester-grade control in the hands of students. Because students are encouraged to revise pieces with which they struggled, assessment becomes part of an ongoing learning process. Students receive thoughtful feedback from the structure of the grading system, while teachers are given a structure to assist them in defining expectations and to communicate those expectations to students.

REFERENCE

National Council Teachers of Mathematics (NCTM). *Assessment Standards for School Mathematics*. Reston, Va.: NCTM, 1995. 