Smith, Margaret Schwan, and Mary Kay Stein. "Selecting and Creating Mathematical Tasks: From Research to Practice." *Mathematics Teaching in the Middle School* 3 (February 1998): 344–50.

Selecting and Creating Mathematical Tasks: From Research to Practice

MARGARET SCHWAN SMITH
AND MARY KAY STEIN

WHAT FEATURES OF A MATHEMATICS classroom really make a difference in how students come to view mathematics and what they ultimately learn? Is it whether students are working in small groups? Is it whether students are using manipulatives? Is it the nature of the mathematical tasks that are given to students? Research conducted in the QUASAR project, a five-year study of mathematics education reform in urban middle schools (Silver and Stein 1996), offers some insight into these questions. From 1990 through 1995, data were collected about many aspects of reform teaching, including the use of small groups; the tools that were available for student use, for example, manipulatives and calculators; and the nature of the mathematics tasks. A major finding of this research to date, as described in the article by Stein and Smith in the January 1998 issue of *Mathematics Teaching in the Middle School,* is that the highest learning gains on a mathematics-performance assessment were related to the extent to which tasks were set up and implemented in ways that engaged students in high levels of cognitive thinking and reasoning (Stein and Lane 1996). This finding supports the position that the nature of the tasks to which students are exposed determines what students learn (NCTM 1991), and it also leads to many questions that should be considered by middle school teachers.

In particular, results from Stein and Lane (1996) suggest the importance of starting with high-level, cognitively complex tasks if the ultimate goal is to have students develop the capacity to think, reason, and problem solve. As was noted in our earlier discussion of Ron Castleman (Stein and Smith 1998), selecting and setting up a high-level task well does not guarantee students' engagement at a high level. Starting with a good task does, however, appear to be a necessary condition, since low-level tasks almost never result in high-level engagement. In this article, we focus on the selection and creation of mathematical tasks, drawing on QUASAR's research on mathematical tasks and on our own experiences with teachers and teacher educators.

MARGARET SMITH, *Pennsylvania State University, University Park, PA 16802, is interested in supporting and studying the professional development of teachers.* MARY KAY STEIN, *University of Pittsburgh, Pittsburgh, PA 15260, studies the professional development of teachers and teaching and learning as they occur in classroom settings.*

"Reflections on Practice" is a look back at the end result of putting theory into practice. For those interested in submitting manuscripts that concern this theme, please send them to editor SUSAN FRIEL, *Mathematics Department, University of North Carolina—Chapel Hill, Chapel Hill, NC 27599-3500.*

Knowing a Good Task When You See One

WHEN CLASSIFYING A MATHEMATICAL TASK AS "GOOD," that is, as having the potential to engage students in high-level thinking, we first consider the students—their age, grade level, prior knowledge and experiences—and the norms and expectations for work in their classroom. Consider, for example, a task in which students are asked to add five two-digit numbers and explain the process they used. For a fifth- or sixth-grade student who has access to a calculator, the addition algorithm, or both, and

Reflection:
Can you think of a task you used that was harder or easier for students than you had anticipated? What factors do you think contributed to the level of difficulty of the task for your students?

for whom "explain the process" means "tell how you did it," the task could be considered routine. If, however, the task is given to a second grader who has just started work with two-digit numbers, who has base-ten pieces available, and for whom "explain the process" means "you need to explain your thinking," the task may indeed be high level. Therefore, when a teacher selects a task for use in a classroom setting, all these factors need to be considered to determine the extent to which a task is likely to afford an appropriate level of challenge for her or his students.

A second step we use in classifying tasks as good is to consider the four categories of cognitive demand described in Stein and Smith (1998):

- Memorization
- Procedures without connections to concepts or meaning
- Procedures with connections to concepts and meaning
- Doing mathematics

Using these categories as templates, we ask ourselves what kind of thinking a task will demand of the students. Tasks that ask students to perform a memorized procedure in a routine manner lead to one level of thinking; tasks that ask students to think conceptually lead to a very different set of thinking processes.

Reflection:
*Consider the eight tasks shown in **figure 1**. How would your students go about solving these tasks? Using the four categories of cognitive demand, how would you categorize each of the tasks for your students?*

In our work with teachers, we have found that they do not always agree with one another—or with us—on how tasks should be categorized. For example, some have categorized task D (shown in **fig. 1**) as a high-level task because it says that students must "explain the process you used" or because it is a word problem. Similarly, some have thought that task F (shown in **fig. 1**) was high level because it used manipulatives and featured a diagram. But we have classified both tasks as low level because each required the use of a procedure as *stated* (task F) or as *implied* by the problem (task D). Neither task presented any ambiguity about what needed to be done or how to do it or had any connection to meaning. So even though the problem might look high level, an observer must move beyond its surface features to consider the kind of thinking it requires.

A Tool for Analyzing Cognitive Demands

ON THE BASIS OF THE FINDINGS REGARDING THE importance of using cognitively demanding tasks in classroom instruction, we, along with our colleague and collaborator, Marjorie Henningsen, created a task-sort activity and a task-analysis guide for use in professional-development sessions to help teachers with the selection and creation of tasks. The task-sort activity consists of twenty carefully selected instructional tasks that represent the four categories of cognitive demand for middle school students. The eight tasks shown in **figure 1** are a subset of the tasks that are included in the sort.

Reflection:
Can you think of other factors that might make a task appear to be high level on the surface but that actually only require recall of memorized information or procedures?

In addition to differing with respect to cognitive demand, the tasks in this activity also differ with respect to other features that are often associated with reform-oriented instructional tasks (NCTM 1991; Stein, Grover, and Henningsen 1996). For example, some tasks require an explanation or description (e.g., tasks A, C, D, and G); can be solved using manipulatives (e.g., tasks A, E, and F); have real-world contexts (e.g., B, C, and D); involve multiple steps, actions, or judgments (e.g., A, B, C, D, E, and G); and make use of diagrams (e.g., A, E, F, and G). Varying tasks with respect to these features across categories of cognitive demand requires an analysis of the task that goes beyond superficial features to focus on the kind of thinking in which students must engage to complete the tasks.

The task-analysis guide **(fig. 2)** consists of a listing of the characteristics of tasks at each level of cognitive demand. It serves as a judgment template—a kind of scoring rubric—that can be applied to all kinds of mathematical tasks, permitting a rating of the tasks. Also included in the task-analysis guide is an example of a task at each level, as shown in **figure 3.** Note that each of the four tasks shown in **figure 3** involves fraction multiplication, yet the tasks vary with respect to the demands they place on students.

Reflection:
How might you use this tool in professional-development sessions to stimulate rich and lively discussions about mathematical tasks and the levels of thinking required to solve them?

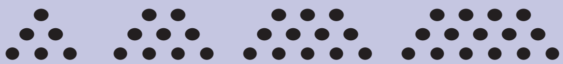
Using the Tool to Facilitate Discussion

TO DATE, THE TASK-SORT ACTIVITY AND THE TASK-analysis guide have been used in a range of settings with preservice and in-service teachers and with teacher educators. In one situation, thirty-three preservice teachers were

TASK A

Manipulatives/Tools: Counters

For homework Mark's teacher asked him to look at the pattern below and draw the figure that should come next.



Mark does not know how to find the next figure.

A. Draw the next figure for Mark.

B. Write a description for Mark telling him how you knew which figure comes next.

(QUASAR Project—QUASAR Cognitive Assessment Instrument—Release Task)

TASK B

Manipulatives/Tools: None

Part A: After the first two games of the season, the best player on the girls' basketball team had made 12 out of 20 free throws. The best player on the boys' basketball team had made 14 out of 25 free throws. Which player had made the greater percent of free throws?

Part B: The "better" player had to sit out the third game because of an injury. How many baskets, out of an additional 10 free-throw "tries," would the other player need to make to take the lead in terms of greatest percentage of free throws?

(Adapted from *Investigating Mathematics* [New York: Glencoe Macmillan/McGraw-Hill, 1994])

TASK C

Manipulatives/Tools: Calculator

Your school's science club has decided to do a special project on nature photography. They decided to take a few more than 300 outdoor photos in a variety of natural settings and in all different types of weather. They want to choose some of the best photographs and enter the state nature photography contest. The club was thinking of buying a 35 mm camera, but one member suggested that it might be better to buy disposable cameras instead. The regular camera with autofocus and automatic light meter would cost about \$40.00, and film would cost \$3.98 for 24 exposures and \$5.95 for 36 exposures. The disposable cameras could be purchased in packs of three for \$20.00, with two of the three taking 24 pictures and the third one taking 27 pictures. Single disposables could be purchased for \$8.95. The club officers have to decide which would be the better option and justify their decisions to the club advisor. Do you think that they should purchase the regular camera or the disposable cameras? Write a justification that clearly explains your reasoning.

TASK D

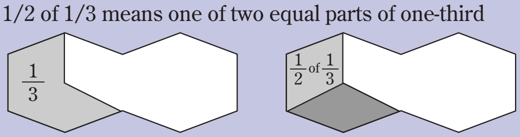
Manipulatives/Tools: None

The cost of a sweater at a department store was \$45. At the store's "day and night" sale it was marked 30 percent off the original price. What was the price of the sweater during the sale? Explain the process you used to find the sale price.

TASK E

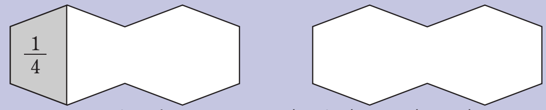
Manipulatives/Tools: Pattern blocks

1/2 of 1/3 means one of two equal parts of one-third



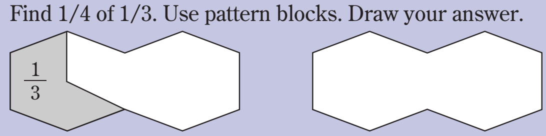
one-third   1/2 of 1/3, or 1/2 × 1/3, = 1/6

Find 1/3 of 1/4. Use pattern blocks. Draw your answer.



one-fourth   1/3 of 1/4, or 1/3 × 1/4, = _______.

Find 1/4 of 1/3. Use pattern blocks. Draw your answer.



one-third   1/4 of 1/3, or 1/4 × 1/3, = _______.

TASK F

Manipulatives/Tools: Square pattern tiles

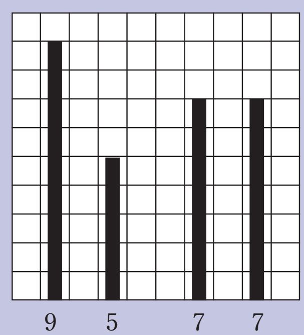
Using the side of a square pattern tile as a measure, find the perimeter of, or distance around, each train in the pattern-block figure shown.

Train 1   Train 2   Train 3

TASK G

Manipulatives/Tools: Grid paper

The pairs of numbers in (a)–(d) represent the heights of stacks of cubes to be leveled off. On grid paper, sketch the front views of the columns of cubes with these heights before and after they are leveled off. Write a statement under the sketches that explains how your method of leveling off is related to finding the average of the two numbers.



(a) 14 and 8   (b) 16 and 7   (c) 7 and 12   (d) 13 and 15

By taking two blocks off the first stack and giving them to the second stack, I've made the two stacks the same. So the total number of cubes is now distributed into two columns of equal height. And that is what average means.

(Taken from Bennett and Foreman [1989/1991])

TASK H

Manipulatives/Tools: None

Give the fraction and percent for each decimal.

0.20 = _______ = _______.
0.25 = _______ = _______.
0.33 = _______ = _______.
0.50 = _______ = _______.
0.66 = _______ = _______.
0.75 = _______ = _______.

Fig. 1 Sample tasks from the task-sort activity

asked to place each of the twenty tasks into one of the four categories of cognitive demand, without the aid of the list of characteristics in **figure 2**. Thus, the teachers were not only sorting but also engaging in discourse about students' levels of thinking as they negotiated definitions for the categories. Once each group had accomplished its assignment, the classifications were tallied in a table. The tally revealed that several tasks had complete or near consensus! Many of these tasks had the hallmarks of a particular category of cognitive demand. For example, task E was categorized by all groups as *procedures with connections*. The discussion brought out the facts that the task focused on what it means to take a fraction of a fraction, as opposed to using an algorithm, such as "multiply the numerators and multiply the denominators"; and that it could not be completed without effort, that is, students needed to think about what their actions meant as they worked on the problem. From the specifics of the example, we began to extract characteristics of the category more generally. In this example, the tasks categorized as *procedures with connections* focus on meaning, require effort, and involve a procedure. Similar discussions surrounding consensus tasks served to develop descriptors of the other categories of cognitive demand.

Reflection:
What do the classifications "procedures with connections" and "doing mathematics" mean to you? How are they alike? How are they different? In what ways can these classifications be helpful in selecting and creating worthwhile mathematics tasks for use in your own classroom?

For other tasks, such as task A, little agreement occurred. Some teachers classified task A as *procedures without connections*; some, as *procedures with connections*; and others, as *doing mathematics*. The ensuing discussion highlighted the fact that no procedure or pathway was stated or implied for task A, yet the group had included the use of a procedure as a hallmark of tasks that were classified as *procedures without connections* and *procedures with connections*. A more focused look at the characteristics of *doing mathematics* brought out the fact that tasks in this category required students to explore and understand the nature of relationships—a necessary step in extending and describing the pattern in task A. The discussion concluded with the preservice teachers deciding to classify the task as *doing mathematics*. By using the established descriptions created by the group as a template against which to judge little-consensus tasks, the group had a principled basis for the discussions it made.

Once teachers had fairly refined ideas of the characteristics of each category of cognitive demand, it was time to start digging deeper. We began to discuss tasks for which they disagreed on the category, for example, *procedures with connections* versus *doing mathematics*, but, for the most part, we agreed on the level of thinking required, for example, high level. We saw an almost even split in terms of the categorization of task G as either *procedures with connections* or *doing mathematics*. After reviewing the criteria established by the group for these two categories, the teachers determined that *procedures with connections* was a better choice, since a procedure was given—leveling off stacks of cubes—and the procedure was connected to the meaning of *average*. The discussion focused attention on the various forms that procedures can take, such as algorithms and general pathways through the problem, and on an important characteristic of *doing mathematics* tasks that this particular task did not possess: the need for students to impose their own structure and procedure.

Reflection:
What other issues might be important to raise in a discussion of tasks? What task would you add to the sort to stimulate additional discussion?

We concluded the session by distributing the task-analysis guide and comparing the teachers' descriptors with those that appeared in the guide. By distributing the guide *after* the task-sorting activity was completed, we did not constrain the earlier discussion by the characteristics listed in the guide and participants had the opportunity to construct a listing in their own language. The long-term goal of this activity was twofold: to raise awareness of how mathematical tasks differ with respect to the levels of cognitive engagement that they demand from students and to facilitate teachers' development of a deep and sustained appreciation for the principles of task selection and design.

Sharing Your Reflections

IN THIS ARTICLE WE SHARED OUR FINDINGS CONCERNING the importance of beginning with a task that has the potential to engage students at a high level if your goal is to increase students' ability to think and reason. The point is that the task you select and evaluate should match your goals for student learning. We encourage you to *(a)* reflect on the extent to which the tasks you use match your goals for student learning, *(b)* reflect on the extent to which your students have the opportunity to engage in tasks that require thinking and reasoning, *(c)* use the eight tasks that are listed in **figure 1** in a discussion with your colleagues, and *(d)* share the results of your experiences through the "Teacher to Teacher" department in this journal.

Levels of Demands

Lower-level demands (memorization):

- Involve either reproducing previously learned facts, rules, formulas, or definitions or committing facts, rules, formulas or definitions to memory
- Cannot be solved using procedures because a procedure does not exist or because the time frame in which the task is being completed is too short to use a procedure
- Are not ambiguous. Such tasks involve the exact reproduction of previously seen material, and what is to be reproduced is clearly and directly stated.
- Have no connection to the concepts or meaning that underlie the facts, rules, formulas, or definitions being learned or reproduced

Lower-level demands (procedures without connections):

- Are algorithmic. Use of the procedure either is specifically called for or is evident from prior instruction, experience, or placement of the task.
- Require limited cognitive demand for successful completion. Little ambiguity exists about what needs to be done and how to do it.
- Have no connection to the concepts or meaning that underlie the procedure being used
- Are focused on producing correct answers instead of on developing mathematical understanding
- Require no explanations or explanations that focus solely on describing the procedure that was used

Higher-level demands (procedures with connections):

- Focus students' attention on the use of procedures for the purpose of developing deeper levels of understanding of mathematical concepts and ideas
- Suggest explicitly or implicitly pathways to follow that are broad general procedures that have close connections to underlying conceptual ideas as opposed to narrow algorithms that are opaque with respect to underlying concepts
- Usually are represented in multiple ways, such as visual diagrams, manipulatives, symbols, and problem situations. Making connections among multiple representations helps develop meaning.
- Require some degree of cognitive effort. Although general procedures may be followed, they cannot be followed mindlessly. Students need to engage with conceptual ideas that underlie the procedures to complete the task successfully and that develop understanding.

Higher-level demands (doing mathematics):

- Require complex and nonalgorithmic thinking—a predictable, well-rehearsed approach or pathway is not explicitly suggested by the task, task instructions, or a worked-out example.
- Require students to explore and understand the nature of mathematical concepts, processes, or relationships
- Demand self-monitoring or self-regulation of one's own cognitive processes
- Require students to access relevant knowledge and experiences and make appropriate use of them in working through the task
- Require students to analyze the task and actively examine task constraints that may limit possible solution strategies and solutions
- Require considerable cognitive effort and may involve some level of anxiety for the student because of the unpredictable nature of the solution process required

These characteristics are derived from the work of Doyle on academic tasks (1988) and Resnick on high-level-thinking skills (1987), the *Professional Standards for Teaching Mathematics* (NCTM 1991), and the examination and categorization of hundreds of tasks used in QUASAR classrooms (Stein, Grover, and Henningsen 1996; Stein, Lane, and Silver 1996).

Fig. 2 Characteristics of mathematical instructional tasks

Lower-Level Demands

Memorization

What is the rule for multiplying fractions?

Expected student response:

You multiply the numerator times the numerator and the denominator times the denominator.

or

You multiply the two top numbers and then the two bottom numbers.

Procedures without Connections

Multiply:

$$\frac{2}{3}\times\frac{3}{4}$$

$$\frac{5}{6}\times\frac{7}{8}$$

$$\frac{4}{9}\times\frac{3}{5}$$

Expected student response:

$$\frac{2}{3}\times\frac{3}{4}=\frac{2\times 3}{3\times 4}=\frac{6}{12}$$

$$\frac{5}{6}\times\frac{7}{8}=\frac{5\times 7}{6\times 8}=\frac{35}{48}$$

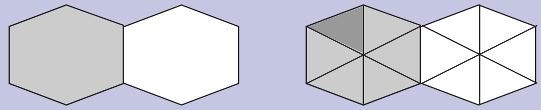
$$\frac{4}{9}\times\frac{3}{5}=\frac{4\times 3}{9\times 5}=\frac{12}{45}$$

Higher-Level Demands

Procedures with Connections

Find 1/6 of 1/2. Use pattern blocks. Draw your answer and explain your solution.

Expected student response:



First you take half of the whole, which would be one hexagon. Then you take one-sixth of that half. So I divided the hexagon into six pieces, which would be six triangles. I only needed one-sixth, so that would be one triangle. Then I needed to figure out what part of the two hexagons one triangle was, and it was 1 out of 12. So 1/6 of 1/2 is 1/12.

Doing Mathematics

Create a real-world situation for the following problem:

$$\frac{2}{3}\times\frac{3}{4}.$$

Solve the problem you have created without using the rule, and explain your solution.

One possible student response:

For lunch Mom gave me three-fourths of a pizza that we ordered. I could only finish two-thirds of what she gave me. How much of the whole pizza did I eat?

I drew a rectangle to show the whole pizza. Then I cut it into fourths and shaded three of them to show the part Mom gave me. Since I only ate two-thirds of what she gave me, that would be only two of the shaded sections.

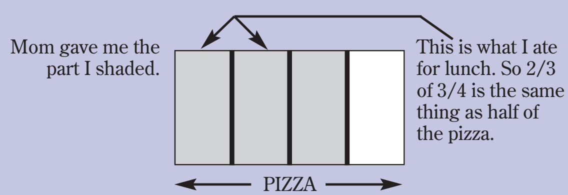


Fig. 3 Examples of tasks at each of the four levels of cognitive demand

References

Bennett, Albert B., and Linda Foreman. *Visual Mathematics Course Guide: Integrated Math Topics and Teaching Strategies for Developing Insights and Concepts*, vol. 1. Salem, Ore.: Math Learning Center, 1989/1991.

Doyle, Walter. "Work in Mathematics Classes: The Context of Students' Thinking during Instruction." *Educational Psychologist* 23 (February 1988): 167–80.

National Council of Teachers of Mathematics (NCTM). *Professional Standards for Teaching Mathematics*. Reston, Va.: NCTM, 1991.

Resnick, Lauren. *Education and Learning to Think*. Washington, D.C.: National Academy Press, 1987.

Silver, Edward A., and Mary K. Stein. "The QUASAR Project: The 'Revolution of the Possible' in Mathematics Instructional Reform in Urban Middle Schools." *Urban Education* 30 (January 1996): 476-521.

Stein, Mary Kay, Barbara W. Grover, and Marjorie Henningsen. "Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms." *American Educational Research Journal* 33 (October 1996): 455–88.

Stein, Mary Kay, and Suzanne Lane. "Instructional Tasks and the Development of Student Capacity to Think and Reason: An Analysis of the Relationship between Teaching and Learning in a Reform Mathematics Project." *Educational Research and Evaluation* 2 (October 1996): 50–80.

Stein, Mary Kay, Suzanne Lane, and Edward Silver. "Classrooms in Which Students Successfully Acquire Mathematical Proficiency: What Are the Critical Features of Teachers' Instructional Practice?" Paper presented at the annual meeting of the American Educational Research Association, New York, April 1996.

Stein, Mary Kay, and Margaret S. Smith. "Mathematical Tasks as a Framework for Reflection." *Mathematics Teaching in the Middle School* 3 (January 1998): 268–75.

The preparation of this paper was supported by a grant from the Ford Foundation (grant no. 890-0572) for the QUASAR Project. Any opinions expressed herein are those of the authors and do not necessarily represent the views of the Ford Foundation. This paper grew out of a research report by Mary Kay Stein, Barbara Grover, and Marjorie Henningsen (1996). The authors wish to acknowledge the helpful comments of Judith Zawojewski on an earlier draft of this article.