Assessing Mathematical Processes: The English Experience

Malcolm Swan

Historically, assessment has been concerned with measuring how completely students have mastered knowledge and skills. More recently, however, increasing attention is being placed on how effectively students can tackle unstructured problems and investigate novel, open-ended situations. Here, the focus of assessment is on how well students have acquired those processes or strategies that guide the choice of appropriate skills and enable students to explore unfamiliar situations. In short, we are becoming more aware of the need to assess how well students perform as mathematicians in addition to how well they have learned mathematics.

Whereas the knowledge of facts and skills can be assessed through short, closed questions, the existence of strategic skills can be assessed only through more open tasks that require students to make choices, to reason and to explain. Traditional forms of timed, written examinations do not allow sufficient time for students to pursue their own lines of inquiry, meet dead ends, plan approaches and so on. A system is needed that allows students adequate time to tackle extended tasks, collaborate with peers, reflect and redraft ideas and polish products. The system also needs to be manageable and rigorous so that the assessments are made efficiently and reliably. For the system to have status, we must also ensure that different assessors use common standards that are externally validated. This article describes such a system, warts and all, that is currently being used throughout England.

Specifying an Assessment System

Traditionally in England, assessment has been norm-referenced. At the end of a period of assessment, all we could say about students was how well they performed in comparison with their peers. We were unable to say anything objective about what the student was able to do. It was impossible to know whether, in the long term, standards were rising or falling. For these reasons, the government decided to introduce a set of national criteria against which assessments would be made (DFE/WO 1991). Mathematics currently has five attainment targets: Using and Applying Mathematics, Number, Algebra, Shape and Space, and Data Handling. The last four targets list the facts, skills and concepts to be assessed in traditional content areas (for example, "can find a fraction of a quantity"); the first lists the mathematical processes to be assessed (for example, "can make generalizations"). Having a separate list for processes has the advantage that strategic skills are given more status than if they are merely absorbed into content lists, but the danger also exists that they are taught and assessed separately from "real mathematics."

Each attainment target has a 10-level hierarchical description. Teachers are required to use these levels to describe how students' concepts and skills develop through their school careers. The levels are not age related. In any mathematics class can be found a range of levels of performance in any individual attainment target. Similarly, an individual student will perform differently across different attainment targets. Teachers are required to keep careful records documenting each student's profile and progress. For students at the age of 7, 11, 14 and 16, the results of these assessments, together with the results of externally supplied national tests, are made public.

The laudable desire to pinpoint precisely what students know and can do has unfortunately led to a proliferation of criteria; the system has become so unwieldy that it is currently undergoing simplification (SCAA 1994). A second problem is that the original intention of the levels—monitoring performance outcomes—has often been misinterpreted, by teachers and textbook writers, as prescribing the order in which mathematics must be taught. This misconception has led to whole classes being taught level 5 before level 6, making nonsense of the differentiation in ability that exists within a class.

Because this article is concerned with assessing mathematical processes, attention is focused on just the first attainment target, Using and Applying Mathematics, which is subdivided into three strands for the purpose of assessment. These strands are (1) applications, (2) communication and (3) reasoning, logic and proof. The strands are described more fully in Table 1. In the original documents, examples are attached to each statement to aid their interpretation.

Table 1. **Attainment Target 1: Using and Applying Mathematics**

Level	Strand 1: Applications	Strand 2: Communication	Strand 3: Reasoning, Logic and Proof
10	Explore independently and constructively a new area of mathematics.	Apply mathematical language and symbolism confidently when handling abstract concepts. Present logical and concise accounts of work resulting from an independent exploration of a new area of mathematics, commenting on alternative solutions.	Handle abstract concepts of proof and definition when exploring independently a familiar or new area of mathematics.
9	Coordinate a number of features or variables of solving problems.	Use mathematical language and symbolism effectively when presenting logical accounts of work. Produce concise justifications of their solutions to complex problems.	Justify their solutions to problems involving a number of features or variables.
8	Make reasoned choices when exploring a mathematical task.	Use mathematical language and symbolism effectively when presenting logical accounts of work, stating reasons for choices made.	Understand the role of counterexamples in disproving generalizations or hypotheses.
7	Follow new lines of inquiry when investigating within mathematics itself or when using mathematics to solve a real-life problem.	Use appropriate mathematical language and notation when solving real-life problems or commenting on generalizations or solutions.	Examine and comment constructively on generalizations or solutions.
6	Pose their own questions or design a task in a given context.	Examine critically the mathematical presentation of information.	Make a generalization giving some degrees of justification.
5	Carry through a task by breaking it down into smaller, more manageable tasks.	Interpret information presented in a variety of mathematical forms.	Make a generalization and test it.
4	Identify and obtain information necessary to solve problems.	Interpret situations mathematically, using appropriate symbols of diagrams.	Give some justification for their solutions to problems. Make generalizations.
3	Find ways of overcoming difficulties when solving problems.	Use or interpret appropriate mathematical aspects of everyday language in a precise way. Present results in a clear and organized way.	Investigate general statements by trying out some examples.
2	Select materials and the mathematics to use for a practical task.	Talk about work or ask questions using appropriate mathematical language.	Respond appropriately to the question "What would happen if . . .?"
1	Use mathematics as an integral part of practical classroom tasks.	Talk about their own work and respond to questions.	Make predictions on the basis of experience.

Assessing Mathematical Strategies

The way in which students are taught from age 11 is deeply affected by the style of the examination that they will eventually face at age 16. In England, probably even more so than in the United States, the examination system drives the curriculum through its backwash effect. How the assessment of the Using and Applying Mathematics strand is carried out by different regional examination boards is presented next.

Teachers currently have a choice, depending on the examination board for which they opt. They may assess students through

- extended tasks that are prescribed by the examination board;
- extended tasks of their own choice but that are bounded within prescribed themes, such as one statistical study, one pure investigation or one with real-life practical applications;
- their own selection of evidence from students' portfolios of work; and
- an externally set examination paper consisting of a number of shorter questions.

This final method is currently proving the most controversial. It is hard to see how such processes as breaking tasks down into manageable steps, posing questions or exploring independently can be assessed in traditional examination formats, and every attempt made to date has proved unsatisfactory.

This article focuses on the use of more extended tasks that are completed in a normal classroom-working atmosphere. Suitable tasks are open-ended, are accessible to the least able students and also afford opportunities to stretch the more able students.

In most situations, teachers are expected to perform the assessments. The job of ensuring the consistency of standards falls on a system called "moderation." Within individual schools, teachers hold standardization meetings at which they award scores to the students' portfolios and rank them in order of merit. Samples of these portfolios are then required by the examination board for inspection by area moderators. Their job is to ensure that different schools are applying similar standards. They will not attempt to alter the ranking of students within a school, but they may adjust all the scores to bring them into line with those from other schools.

A Sample Task

The following activity was sent to all schools to exemplify the type of task that supports the assessment of strategic skills. See Figure 1. Students are given octagonal tiles made from cardboard and work on the problem in groups. They may decide to investigate the number of tiles in each loop or the number of free edges inside, outside or in total. This initial investigation can then be extended according to the capabilities of students. For example, pupils might focus on the following concepts:

- The different loops that can be made from a specific number of tiles
- The shapes inside loops and the shapes made by joining the centre of each tile (see Figure 2)
- The ways of recording shapes (Figure 3)
- The symmetrical properties of the tiles and loops
- Some ways to change the rules for making the loops
- The angle sum of the shapes inside the loops
- Whether all regular polygons can make rings (Figure 4)
- How to prove that an odd number of octagon tiles will not make a loop
- Using Logo to draw the shapes obtained

Assessing the Work

Figure 5 is part of a response from one student, Alison. She starts by sketching the inside shapes made by loops of four, six, and eight octagons and finds that no loops exist for three, five or seven octagons. She then looks for patterns in the number of inside and outside free edges. The reader might try to assess this response with the "Using and Applying Mathematics" attainment target (Table 1) before reading on.

To assess this work, teachers match aspects of this performance to the general criteria in each strand. This task is not always straightforward.

Consider strand 1, Applications. Alison has thought about the shapes that she can make with three, four, five, six, seven and eight octagons in turn. She has thus broken down the investigation into stages. This process could be said to show evidence that she can "carry through a task by breaking it down into smaller, more manageable tasks" (level 5). She has clearly not posed her own questions or followed new lines of inquiry, so no evidence demonstrates that she has reached levels 6 or 7.

Look at strand 2, Communication. Her work uses diagrams, tables and text. Perhaps this presentation suggests that she can "interpret situations mathematically, using appropriate symbols or diagrams" (level 4) or "interpret information in a variety of mathematical forms" (level 5). Although hard to judge, level 5 perhaps requires a wider variety of forms than that shown here, such as graphical and symbolic forms. Alison has not yet shown that she can examine critically the mathematical presentation of information (level 6). This ability could perhaps be shown if she is asked to criticize the presentation of her work and invited to redraft it.

Next examine strand 3, Reasoning. Her conclusion that "you multiply the number of octagons by 6" to get the total number of free edges shows that Alison can "make a generalization and test it" (level 5). She has not explained why this pattern occurs—for each tile in a loop, two edges are touching and six are free—so she has not yet shown that she can "give some degree of justification" (level 6).

This analysis leads to several important points:

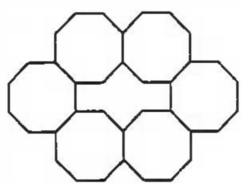
- Matching specific performances to general criteria is a subjective business. A wide variety of interpretations may be given to each level's descriptor. In practice, however, teachers are given examples of assessed students' work to guide their judgments, and the process of consultation helps teachers reach a consensus "feeling" for what a level's descriptor means. For situations in which externally supplied tasks are used, task-specific descriptors are often given.
- The demand, or level, of a process criterion is not meaningful unless it is related to a particular context. Finding a generalization of the number of inside free edges, $3n - 8$, is much harder than finding a generalization for the total number of free edges, $6n$, where n is the number of tiles used. Thus a level cannot really be attached to "Make a generalization and test it" unless the context is specified more closely. Again, assessment falls back on the judgments of teachers.
- Even if levels were well defined, assigning a level to a student on the basis of a single fragment of evidence is clearly not possible. In addition, it is unclear whether the "best" or a "typical" performance of the student should be assessed. One examination board, for example, specified that the teacher select "two pieces of evidence which represent the best sustainable work of the student in each strand."
- If students are unaware of which aspects of performance are being assessed, they are unlikely to display these aspects. In the foregoing example, we cannot say whether Alison was able to extend the problem or use a graph. She was not asked to do these things. One alternative is to introduce new scaffolded prompts, but this addition destroys the openness of the task. A second alternative is to make students more aware of the criteria on which they will be judged. Some teachers have developed student-friendly versions of the criteria or even samples of assessed work for students to discuss.
- Most tasks do not permit students to display performance at the full range of levels. In the foregoing task, students had no opportunity to reach the highest levels in "communication," for example. This limitation means, therefore, that a carefully balanced range of tasks must be offered.

Figure 1. **A Task Supporting the Assessment of Strategic Skills. Reproduced with permission of SCAA (1992).**

Octagon loops

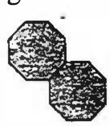
Making the loops

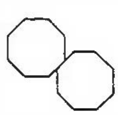
This is an octagon loop. It is made by joining octagon tiles together. There is only one space in the middle of the loop.

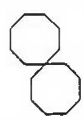


Rules for joining the tiles

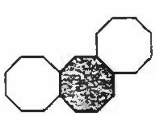
Line up the edges exactly. Not like this . . . or like this . . .







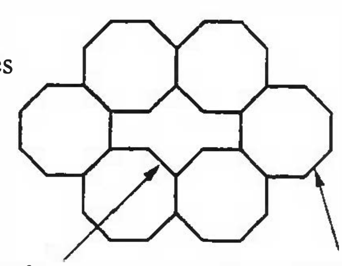
Make sure that each tile only touches two others. Not like this



Make some loops. Use any number of tiles up to 10. Try 5 different loops.

Looking at number patterns

6 tiles



10 inside free edges 26 outside free edges

Total number of free edges 36

Can you tell how many free edges there will be in your next loop before you make it? Investigate

Concluding Comments

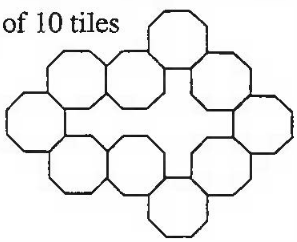
So What Is All This Assessment For?

Ongoing assessment is intended to be formative, whenever individual profiles are kept of each student's progress. In many situations, these profiles have ensured more continuity when students change schools or move between classes. Many teachers, however, see all this record keeping as an unnecessary form of bureaucracy that adds nothing of significance to what they already know about their students. Many have simply refused to participate in the assessment.

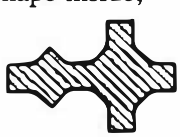
In some schools, especially those in which students are involved in their own assessment, these profiles can serve to inform and motivate students. In one school, for example, each student is given a list of specific content-and-process learning objectives at the start of each extended mathematics activity. These objectives are taken directly from the National Curriculum Framework and are translated into simple, task-specific English. At the end of the activity, students assess their own performance against these targets. The teacher then assesses the work and discusses any discrepancies with each student privately. Assessed portfolios of the students' best work are built. Over a period of time, students thus become much

Figure 2. **Loops, Octagons and Resulting Shapes**

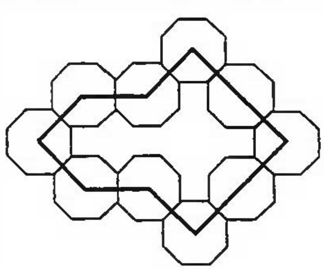
This loop of 10 tiles



produces this shape inside;



when the centres of the octagons are joined,



this shape results.

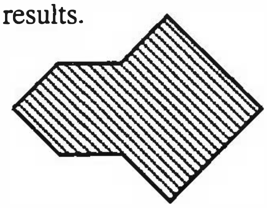


Figure 3. **Encoding the Perimeters**

The perimeter of the shape inside this loop may be described by this code: 1, 2, 4, 1, 3, 1, 3, 1, 4, 2. It could also be described by its area.

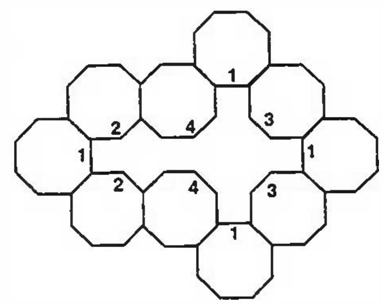


Figure 4. **Students Are Asked Whether Regular Polygons Can Make Rings**

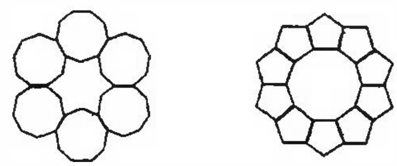
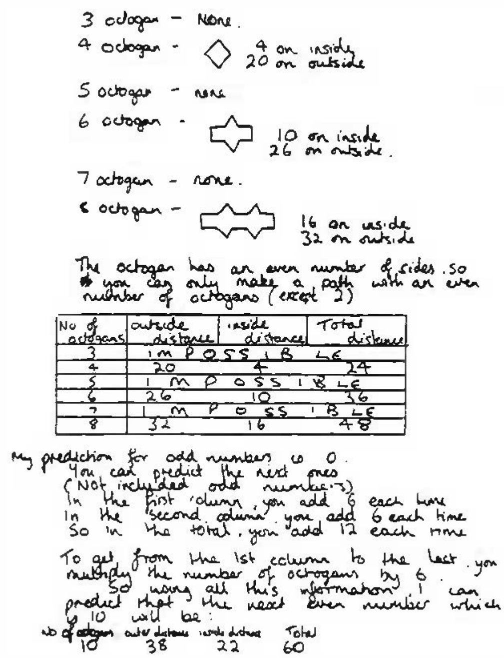


Figure 5. **Alison's Response to the Task**



more aware of their own progress. One noticeable outcome is that students begin to realize that strategic skills are legitimate goals for learning.

Unfortunately, when assessments are made for summative purposes, levels for process get added, weighted and combined with content scores to give relatively meaningless lettered grades. We still have "grade A" students and "grade G" students. Norm-referencing still plays an important part in checking that grade boundaries are "accurately" placed, since populations are not expected to change significantly from year to year. This tendency to reduce people to numbers or letters for selection purposes is unjustifiable. Any system of assessment can be turned from one that celebrates positive achievements into one that just placed crude, meaningless labels on people.

References

Department for Education and the Welsh Office (DFE/WO). *Mathematics in the National Curriculum.* London, U.K.:HMSO, 1991.

School Curriculum and Assessment Authority (SCAA). *Key Stage 3 School Assessment Folder (Part Three) Materials to Support the Assessment of Using and Applying Mathematics.* London, U.K.: Author, 1992.

Mathematics in the National Curriculum. Draft proposals. London, U.K.: HMSO, 1994.

Reprinted with permission from Mathematics Teaching in the Middle School, *Volume 1, Number 9, March-April 1996, pp. 706–10. Minor changes have been made to spelling and punctuation to fit ATA style.*

What is the solution of the following cryptarithm puzzle?

```
 MIX
 FUN
+AND
MATH
```