

---

---

# Performance Assessment: The Message from England

---

*Although many educators in the United Kingdom think performance assessment benefits students, critics are challenging the rigor of the evaluations and advocating dropping practical tasks from examinations.*

DESMOND L. NUTTALL

All the world has become interested in performance assessment. It is now agreed, even in the U.S.A., the home of multiple-choice tests, that such tests do not measure all the important aspects of learning, especially higher-order thinking and practical skills. In England, where performance assessment is not new, there is little doubt that this form of evaluation can improve teaching and learning. Nevertheless, many practical tasks are being dropped from national examinations to simplify the process and save money. Let's look back in time to understand why the trend toward authentic forms of assessment went into reverse.

## **The Assessment of Performance Unit**

England's Assessment of Performance Unit (APU), established in the mid-1970s, fulfilled a similar function in England to that of the National Assessment of Educational Progress in the United States: it provided a light monitoring of educational attainment over time. Early NAEP surveys, like some early state surveys (for example, Michigan's), assessed a wide range of tasks, including practical activities, but subsequently dropped them mainly on financial grounds.

The history of the APU has been somewhat different. The surveys used two methods: traditional assessment and practical assessment. The practical surveys, which were very imaginative, especially in science and mathematics, required the assistance of

visiting assessors to test children one at a time, or in pairs. Although many had feared that the sampling of the APU would rapidly evolve into a method of comparing the performance of different school districts or even of different schools, in practice, teachers welcomed it. They valued the work of the APU for its twin effects of describing student achievement on a much broader front than was typically tested and for providing a mine of ideas for classroom assessment.

## **Public Examinations**

Meanwhile, the traditional public examinations taken by most of the population in Grade 11 and by the elite (around 20 percent of the population) in Grade 13, remained in much the same old format. At Grade 11, there were two examinations, one for the highest attaining students and the other for the rest of the population (mirroring the old division between selective grammar schools and the rest). As comprehensive schooling gained ground, the dual examination system lost its basic rationale, and by 1970 there were calls for a merger.

After a major program of feasibility and development studies, Her Majesty's Inspectorate of Schools, a group with high prestige and much influence in education, recommended the establishment of a common system of examining, including more performance assessment. The Inspectorate's reasons are not hard to find. In a national survey of secondary schooling, published in 1979, the

group made a damning indictment of teaching and learning in the two years leading up to the old Grade 11 examinations:

... the effect of the dominating pursuit of examination results was to narrow learning opportunities. . . . In at least one-fifth of the schools the demands of public examinations appeared to be an important factor in the impoverishment of reading. . . . In many subjects writing tended to be stereotyped and voluminous. . . . In mathematics, examinations were one factor in encouraging narrow, repetitive practice and standard routines divorced from applications to real situations. . . . Similarly in science, the emphasis on content and concepts unrelated to the pupil's experience . . . was precluding investigational activity even where examinations required assessment of such skills.

The Government, never hasty in implementing other people's ideas, took eight years to agree to establish a new examination, the General Certificate of Secondary Education (GCSE).

Every subject in the GCSE contains some element of performance assessment: investigational work in mathematics, oral work in English and in modern languages, portfolios in English, and projects (extended writing up of an investigation carried out and usually chosen by the candidate) mostly in the humanities and the social sciences. Most of this work is graded by the candidate's own teacher, subject to monitoring by an examining body. Monitoring usually occurs through visits by a part-time moderator whose own grading standards are extensively monitored by a Chief Moderator (who is also responsible for training the moderators).

In the four years that the GCSE examination has been running, the proportion of students awarded the top three grades (of the seven that can be

achieved) has increased substantially. Most people hail this as proof that educational standards are rising (even allowing for some grade inflation as a result of greater familiarity with the examination on the part of teachers, students, and parents). Some, though, view the increases as evidence that examiners' standards are slipping and that performance assessment is making the examination far too easy. They point out that students (especially middle-class ones) can get considerable help with assignments from their parents and that essays can be polished before they become part of the portfolio.

The majority's view, however, is that GCSE has been an educational success. Learning and teaching have changed to accommodate the new forms of assessment, and there are now many more student-centered activities in Grade 10 and 11 classrooms. Assessment can, it appears, be a device to change (and improve) teaching and learning. This lesson has not gone unnoticed in the United States. Michael Cohen and Lauren Resnick, who direct the New Standards Project, are experimenting with ideas for a national examination system based heavily on performance, which would help to motivate young people to learn and foster a broader range of skills in U.S. classrooms.

#### **National Curriculum Assessment**

The power of assessment to control the curriculum and what goes on in classrooms, and perhaps to gain votes, did not escape the Thatcher Government. Its 1988 Educational Reform Act was one of the most substantial pieces of English legislation on education this century. Among many other things, the act mandated the creation of a National Curriculum and an associated assessment system for all

students in Grades 2, 6, 9, and 11 (in the last case incorporating GCSE). The National Curriculum was clearly designed to place what was taught in the schools under the control of the Government.

Officials, however, paid less attention to how the curriculum was taught and assessed, leaving the design of an

---

## **Every subject in the General Certificate of Secondary Education contains some element of performance assessment.**

---

assessment system to a panel of educationists. The panel followed current good practice, as exemplified in the GCSE examination, and proposed extensive performance assessment. The system was to be an amalgam of the teacher's assessment of the progress of each child (against national standards as defined in the National Curriculum) and the child's performance in a series of Standard Assessment Tasks designed centrally but administered and graded by the child's own teacher. The activities were to consist of ones very like those employed every day in the classroom, so much so that the children might not even realize that they were being assessed; the "Standard" in "Standard Assessment Task" simply indicates that the task is common across England. Small-scale trials were held in 1990, and a fully fledged national pilot (for every Grade 2 child in the

kingdom) was held in the summer of 1991.

One could hardly recognize the 1991 Standard Assessment Task as a test. First, it was sent out weeks before the assessment period to enable teachers to dovetail it into their own teaching plans. Second, no standardized instructions were prescribed; teachers were encouraged to introduce the tasks in their own words. They were to be done either individually or in small groups over a period of at least three weeks and were estimated to occupy about 30 hours of teacher time.

### Examples of Standard Assessment Tasks

Examples from English and science illustrate the range of tasks assessed.<sup>1</sup>

**English.** Teachers are to select stories targeted at the average 7-year-old "that will give rise to an extended discussion." Additional instructions to teachers include:

- Read the story aloud to a small group of children.
- Lead a discussion about the story; encourage children to contribute fully and to think further about their contributions.
- Ask further questions, if necessary, to draw out children's understanding of the structure of the story.
- To be successful in this assessment, children need to understand story structure in terms of the opening, the unfolding events, and the conclusion. Their contributions should be more than just short answers to direct questions.

**Science.** This task, which is aimed at the average 7-year-old, starts with a long list of materials the teacher might need, including containers of water for floating and sinking investigations. The instructions to teachers continue:

- Set out all the apparatus and give each child four objects including: (1)

at least one object that has some features which are easier to see using a hand lens (for example, a twig); (2) at least two objects that will float (for example, a twig and an apple); (3) at least one object that will sink (for example, a stone); (4) at least one heavy object that will float (for example, a piece of fruit).

## Standard Assessment Tasks were to consist of activities so very like those employed every day in the classroom that the children might not even realize they were being assessed.

- Present this part of the activity to the children as a challenge to find out as much as they can about why some objects float and some sink.

- Ask the children to tell you *what* they think will happen if one of their objects, which you name, is placed in a container of water — and *why*. With each child, repeat this with a second object.

- Explain to the children they are to draw, weigh, make a prediction whether each object will float or sink, and then test each object to see whether it will float or sink. On the charts supplied, have them record their predictions, the outcomes of their

tests, and the weight of each object.

- Explain to the children that they should observe and begin to draw each object without using a hand lens and then finish the detail of their drawing of each object with the use of a hand lens.

- Allow the children to carry out their practical work and complete their charts and drawings.

- When the children have completed their activity, ask each child what he or she found out about why some objects floated and some sank.

- Now ask each child three questions to see if he or she can collate information from the chart.

If children meet the standards for the average level, teachers are instructed to have them proceed to the next level (again defined by explicit standards). The next task is for the children to investigate the hypothesis that heavy objects sink.

### Evaluation of the Assessment

In most cases, the children enjoyed the activities on the 1991 Standard Assessment Tasks and were not made anxious by them. Many teachers said that they had learned new things about the children's attainments, adding that they had often not rated the children as high as the assessment did. However, they found the tasks to be demanding of their time and energies; invariably, to prepare, administer, and grade them required an average of 44 hours. In addition, it was almost impossible to administer the Standard Assessment Tasks and look after the rest of the class simultaneously. Most teachers were able to arrange for another teacher to assist with the class, but the interruption of normal education was substantial, as were the extra costs incurred.

Also, as noted earlier, some people have felt that performance assessment is making the examination less

---

---

## Performance assessment can be made to work, and both teachers and students welcome it.

---

rigorous. Unfortunately, the minority who take this view have the ear of Prime Minister John Major, who has decided that the amount of performance assessment must be significantly curtailed. His exact words were: "By testing I do not mean some weird experiment in a corner. [A reference to the Standard Assessment Tasks.] What I mean is pencil-and-paper testing for a classroom so that people have a measure of how they are doing — see if there is a problem so that you can put it right."

As a consequence, the plans for national assessment at Grade 2 in 1992 (and in later years for the other grades) have been changed dramatically. For example, the day after the Prime Minister's statement, the development contracts for Grade 9 Standard Assessment Tasks were canceled, and new tenders were invited for the creation of paper-and-pencil tests to the virtual exclusion of the sort of practical and naturalistic tasks that had been under development. At Grade 2, also, the tasks have been redesigned so that many can be administered to the whole class at the same time, and some of the most time-consuming ones have been dropped. In 1992, the administration of the Standard Assessment Tasks should be much simpler and take up much less time. In addition, teachers will have had ample

time to prepare for it, since it was circulated several months before it had to be given.

### Reporting of Results

One aspect of National Curriculum Assessment that has not yet been put to the test is the reporting of results. A major justification for doing so is to give parents more information about their child's progress, but there are other agendas.

Scores will be made public school by school and classroom by classroom. The publication of the former in local league tables is designed to give information to parents that will help them in choosing a school for their child. The publication of the latter risks a return to the Victorian era, when teachers were paid according to the attainment of their students.

Publication of results in this manner puts much emphasis on comparability of grading between different schools and a move to more objectively scorable tests. The introduction of a more rigorous system of moderation will surely be demanded when scores begin to be published (probably in 1993). Making fair comparisons also requires comparing like schools with like (and like classrooms with like), as is done in California and several other states. The Government is resisting anything but the publication of "raw results," but it is almost certain that school districts will also produce results taking the characteristics of the student body into account.

### Lessons from England

Performance assessment can be made to work, and both teachers and students welcome it. Through the new GCSE examination in England, the process has improved teaching and broadened the range of skills appraised. Performance assessment

demands new skills of the teacher as well, and a new professionalism. Performance assessment can only take off with the will and expertise of the whole profession and with the trust of parents and politicians.

I suggest that the United States might learn two other things from the English experience: first, the cost of

---

## The cost of performance assessment, both financially and in terms of the time of teachers, is immense, and, despite all the care and effort, some will not view it as rigorous enough.

---

performance assessment, both financially and in terms of the time of teachers, is immense; and, second, despite all the care and effort, some will still not view it as rigorous enough. □

*Source: 1991 Key Stage 1 Teacher's Book.* (London: Department of Education and Science/Welsh Office).

**Desmond L. Nuttall** is Director of the Centre for Educational Research, London School of Economics and Political Science, Houghton St., London, England WC2A 2AE, and Professor of Curriculum and Assessment Studies at the University of London Institute of Education, 20 Bedford Way, London, England WC1H 0AL.

Copyright © 1992 by the Association for Supervision and Curriculum Development. All rights reserved.