

Assessment of Mathematical Elements in Engineering and Science Subjects*

SERGEI S. SAZHIN

School of Engineering, Faculty of Science and Engineering, University of Brighton, Cockcroft Building, Brighton BN2 4GJ, UK. E-mail: s.sazhin@bton.ac.uk

The paper compares two methods of assessing students' understanding of the mathematical elements in engineering and science subjects. It is shown that there is a clear overall correlation between the results of a formal assessment (the number of correct answers measured as a percentage) and in-depth assessment (the actual mark awarded to students when the details of their analysis, including explanations and presentation, has been taken into account). This reflects the consistency of both approaches. At the same time the scattering of the actual marks awarded (in-depth results plotted against the formal assessment results) indicates that formal assessment can often be a poor indicator of the students' knowledge. There is a general tendency for students to gain more marks when their work is subjected to an in-depth assessment (instead of a formal assessment). This tendency, however, is reversed if in-depth assessment marks are compared with the results of a widely used formal assessment method based on multiple choice answers. It is recommended that in-depth assessment replaces formal assessment whenever possible.

INTRODUCTION

NON-SCIENTISTS/ENGINEERS often believe that assessing mathematics comes down to ticking and crossing right and wrong answers. In reality if we want to measure the students' understanding and to give a fair grade, we have to consider less tangible evidence, such as the ability to reason a line of thought, and to present knowledge in a coherent way.

In this paper I attempt to give a comparative analysis of two approaches to the assessment of mathematical knowledge. One, based on the number of right and wrong answers, and the other, based on attempting to assess the students' in-depth understanding of the subject. A classical example of the first approach to assessment is the one used in multiple choice examinations. The second approach is the one which I have been using in the assessment of my engineering students' exam scripts. The criteria of my assessment are given in the Appendix. Copies of these criteria are given to students at the beginning of each module so that they understand what is expected of them. The comparison is based on the mathematical elements in engineering and science subjects taught at the School of Engineering of the University of Brighton.

ANALYSIS

Despite the considerable variety of methods of assessment [1–5], they can roughly be divided into two main groups: formal assessment and in-depth

assessment. This division can be compared to the division of learning methods into surface and deep learning [6, 7].

A classical example of formal assessment is a multiple choice examination. This assessment, however, can be used in conventional examinations as well, when the full marks are given for correct answers, and zero marks for the wrong ones. The main advantage of this approach to assessment is that it is probably the least time consuming for an examiner. Also it encourages students to concentrate on the main goal of the test: to obtain the correct numerical answer. This is particularly important for engineering students in view of their future practical work. An engineer can have a brilliant idea of how to design a bridge, but if the calculations turn out to be wrong then the bridge will collapse.

The downside of this approach is that it does not make any distinction between students who fail to obtain the correct answer because of their misunderstanding of the subject and those who made a silly arithmetic mistake in their calculations. On the other hand, in a multiple choice examination students who do not obtain the correct answer can still score by guessing—in the case of the choice of four answers, students on average can get about 25% even if they have no knowledge of the subject at all. Also, it tends to be rather difficult to avoid cheating in this sort of assessment.

The main motivation behind an in-depth assessment then is to provide a balanced assessment of the correctness of the results, students' understanding of the subject and their ability to present the results of their analysis in a coherent and easily understandable way.

This in-depth assessment has clear advantages

* Accepted 23 June 1999.

when compared with the formal assessment, but it has a number of disadvantages. This kind of assessment requires much more of the examiner's time.

Sometimes the criteria of this assessment are not clear and not fully understood by the students (although, in my example, I tried hard to make these criteria as clear as possible: see the Appendix).

In what follows I will attempt to give a qualitative comparison between formal and in-depth assessments based on one of the tests which I administered (core science, foundation year). The test required students to present the detailed solutions with their workings but were marked in two different ways—one according to the correctness of the final answers and the other in-depth assessment according to the criteria in the Appendix. Assessment in other subjects which I teach (engineering science, thermofluids, fuels and combustion, heat transfer, fluid dynamics) show the same tendencies as core science.

The mark which would have been awarded to students if they were formally assessed would have been exactly equal to the percentage of correct answers. (Note that this mark is equal to a multiple of 10%.) The actual mark awarded to students was based on the criteria specified in the Appendix. We believe that these criteria ensure a proper in-depth assessment.

The results of my comparison are shown in Fig. 1. Each block square in this figure represents the mark of an individual student. The

x-coordinate of this square indicates the percentage of correct answers, while the y-coordinate indicates the actual mark awarded to a student. The total number of squares is 63, although some squares coincide and are not distinguishable.

As follows from Fig. 1, there is a clear correlation between both of these marks. This indicates that on average there is no contradiction between in-depth and formal assessments. At the same time there is a considerable scatter of the results around the line:

$$\text{in-depth assessment mark} = \text{percentage of correct answers}$$

This scatter indicates that the marks awarded to individual students based on the formal assessment can considerably exceed or be much less than the marks awarded based on in-depth assessment.

The average actual marks based on in-depth assessment turned out to be slightly higher than those based on the formal assessment. This takes into account the fact that many students obtained wrong answers but demonstrated a certain level of understanding of the subject. For example, a number of students failed to obtain a single correct answer but still scored something in in-depth assessment (although none managed to get a pass mark of 30%). As already mentioned, in a formal assessment based on the multiple choice approach (when a choice of 4 answers is given) students could get on average of 25% without being able to answer any of the questions (much higher than

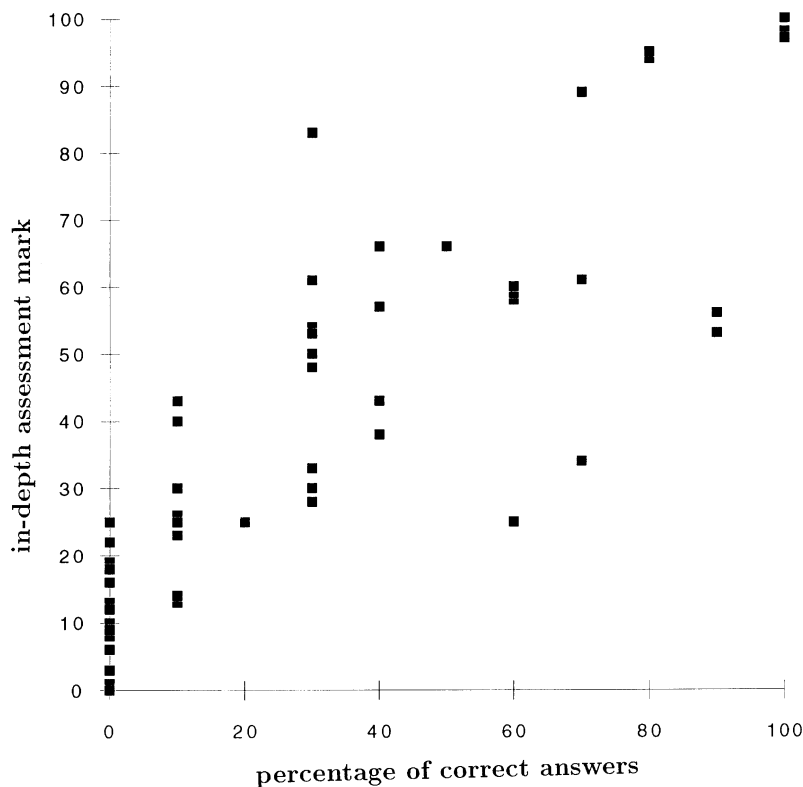


Fig. 1.

the average mark awarded based on in-depth assessment).

Although the number of cases when the mark based on an in-depth assessment was lower than the mark based on the percentage of correct answers was relatively small, they are significant in the appreciation of the underlying philosophy behind an in-depth assessment. The situation occurred in the following cases:

- a) The results were correct, but the presentation was poor. If the presentation was so poor that I could not understand the work, then the students got zero marks for particular questions. The mark was reduced when there were excessive corrections in the script, as I believe that these corrections reflect the lack of confidence in the subject. The mark, however, was not reduced for poor handwriting.
- b) The results were correct but not enough explanations were given. If no explanations at all were given then zero mark was awarded irrespective of whether the answer was correct or not. The same took place when the wrong solution was followed by the correct answer. This measure safeguarded the fairness of the test against possible cheating. (The students were repeatedly told that the objective of the test was to check their understanding of the

subject and not whether they were able to get the correct answer.)

CONCLUSIONS

A detailed comparative analysis of formal and in-depth assessments in engineering and science modules is presented. Based on the results of assessments of the foundation year core science module it is concluded that there is a clear overall correlation between the results of the formal assessment (measured as a percentage of correct answers) and in-depth assessment (the actual mark awarded to the students based on criteria outlined in the Appendix), which reflects the consistency of both approaches to assessment. At the same time a considerable scatter of the actual marks awarded based on in-depth assessment relative to the formal assessment indicates that formal assessment can often be a poor indicator of the depth of students' knowledge. There is a general tendency of the in-depth assessment mark to exceed the formal assessment mark. This tendency, however, is reversed if in-depth assessment marks are compared with results of widely used formal assessment based on multiple choices. It is recommended that in-depth assessment replaces formal assessment whenever possible.

REFERENCES

1. S. S. Sazhin, Assessment in Russian Universities, *Higher Education Review*, **26**, (2) (1994) pp.9–16.
2. I. Hextall, Marking work, in *Exploration in the Politics of School Knowledge*, ed. Whitty, G. and Young, M., Nafferton Books, Driffield, (1976) pp.65–74.
3. C. Carter, Higher education for the future, in *Why and How we Examine*, Blackwell, Oxford, ch. 7, (1980) pp.83–92.
4. P. Broadfoot, Alternative to public examinations, in *Assessing Educational Achievements*, ed. D. Nuttall, Lallmer Press, Lewis, (1986) pp.54–77.
5. S. S. Sazhin, Teaching mathematics to engineering students. *Int. J. Eng. Educ.* **14**, (2), (1998) pp.145–152.
6. F. Marton and R. Saljo, Approaches to learning, in *The Experience of Learning*, ed. F. Marton, D. Hounsell and N. Entwistle, (1997) pp.39–58.
7. L. Svensson, Skill in learning and organising knowledge, in *The Experience of Learning*, ed. F. Marton, D. Hounsell and N. Entwistle, (1997) pp.59–71.

APPENDIX

Criteria of assessment

Correct and—well presented—answers to the questions with detailed and correct explanations are rewarded with the maximum mark. Correct means not only the correct final result but also the most rational way to get it. If the criteria for the maximum mark are not satisfied, then the mark is reduced according to the following guidelines.

- **Correctness.** The correct answer does not automatically lead to the maximum mark and the wrong answer does not necessarily lead to a zero mark. The correct answer without any explanations is generally not accepted. At the same time, if a mistake is made at the final stage of an otherwise correct analysis, then the reduction of the final mark can be as small as 20%. The choice of a correct but not the most effective solution leads to a 10% reduction of the final mark.
- **Explanations.** As already mentioned, answers without explanations are not generally accepted. Correct formulae without adequate verbal explanations

could result in the reduction of the final mark by up to 30%, even in the case when the final result is correct.

- **Presentation.** Answers which are presented in a way which interferes with their adequate understanding will not be accepted. Bad presentation of otherwise understandable answers can lead to a reduction of up to 20% in the final mark.

Sergei Sazhin graduated from Physics Department of St Petersburg University (Russia) in 1972, having completed 5.5 years of undergraduate studies. Between 1972 and 1982 worked as a researcher in the Institute of Physics of St Petersburg University. In 1978 was awarded a research degree of the Candidate of Physical and Mathematical Sciences (Russian equivalent of Ph.D.). Between 1982 and September 1988 worked as a private scientist. Between 1988 and September 1996 worked as a research fellow in Physics Department of Sheffield University (UK) and as a research scientist in the computational fluid dynamics company Fluent Europe (UK). Since September 1996 he has been working as a Senior Lecturer in the School of Engineering at Brighton University (UK). Current research interests include computer modelling of combustion processes in internal combustion engines, thermal radiation, dynamics of fuel sprays and gas flow instabilities. Author of over 160 research papers (mostly in international refereed journals) and two monographs: Natural radio emissions in the Earth's magnetosphere (Nauka Publishing House, Moscow, 1982) and Whistler-mode waves in a hot plasma (Cambridge University Press, 1993). Chartered Physicist, Fellow of the Institute of Physics (UK).