Evaluating the Weekly-Assessed Tutorial Sheet approach to assessment: background, pedagogy and impact

Mark Russell
Principal Lecturer
Aerospace, Automotive and Design Engineering
M.B.Russell@herts.ac.uk, ext 5083

Summary
This paper presents the results of using student-unique, weekly-assessed tasks to overcome a low examination performance in a first-year engineering module. In this instance, the weekly assessment tasks were created by off-the-shelf and bespoke software to form an integrated computer-assisted assessment (CAA) programme. This programme set, delivered, collected, marked and provided prompt feedback on the students' work. The CAA was a set of student-unique Weekly-Assessed Tutorial Sheets (WATS). The rationale for the modified assessment strategy is presented together with the examination performance figures before and after the introduction of WATS.

Background

There is a dynamic in higher education that is changing and evolving as a consequence of, *inter alia*, the UK Government's policy on widening participation and institutional desires to consolidate, and seek further, efficiency gains. Teaching teams now, for instance, are faced with increased student numbers and, often, an attendant reduction in face-to-face lecture time. Moreover, some modules within traditional Engineering degree programmes still rely on students entering with a competency in Engineering Mathematics and a foundation in Engineering Science. Although often anecdotal, there is some evidence to suggest that students' mathematical competency may not meet the expectations of the module teaching teams. Issues relating to student difficulties with Engineering Mathematics are not peculiar to the University but are in fact more widespread. This is evidenced by the wealth of conferences dedicated to teaching Engineering Mathematics as well the ever-increasing provision of new and additional support materials. (Green et al, 2003; Williamson and Green, 2004). Within this changing and difficult backdrop, however, there is both the need to maintain an assessment of the learners with a fair and robust system and an increased recognition of the importance of assessment to help the learning process. In many instances CAA has been used to ease the additional burdens that assessment brings to staff whilst still providing a learning-oriented experience. This paper, therefore, presents the background to a CAA project delivered to a core, first-year engineering module – Fluid Mechanics and Thermodynamics. The rationale for the modified assessment strategy is presented along with the underpinning pedagogy and its impact on student examination performance.

The module

The Fluid Mechanics and Thermodynamics module is delivered to around 130 students. Lectures to 130 students are supported by small-group tutorials and two laboratory studies (group size around 25). Learning is enhanced by active use of StudyNet, via a range of out-of-class activities and the provision of lecture notes and additional support materials. During the 2001/02 session the assessment comprised a phase test, delivered towards the end of the semester, two laboratory reports and an end-of-module examination. The culmination of these learning experiences was poor examination performance: 51% of the students scored less than 35% in the end-of-module examination.

The primary issues surrounding the poor student performance are believed to be the learners' lack of engagement with the module's supporting materials and the in-module phase test not actively supporting their learning. The first issue is evidenced by observations at the small-group tutorial sessions, where it became apparent that many students had not attempted the consolidatory/practice questions issued by the teachers, and by noting that the associated discussion forum received questions that were out of sequence with the lecture series (Russell, 2003). Hence opportunities to enhance learning, by building on the knowledge and the students' understanding that should have been acquired earlier during the module, were subsequently reduced.

Secondly, because of its timing, there were concerns over whether the phase test could actively encourage learning and develop the learners' understanding of the subject. With student numbers of 130, the opportunities to provide prompt and hence useful feedback are reduced. Furthermore, the nature of the one-off phase test provided little motivation for the learners to pick up their books outside the formal contact time.

Influences of the appropriate pedagogy

Prior to discussing the revised assessment and its impact on student performance, a review of the associated pedagogy and literature is given.

Many commentators suggest a strong relationship between assessment and learning. This relationship is exemplified by statements such as, 'The curriculum is defined by assessment' (Ramsden, 1992) and, 'what influenced students most was not the teaching but the assessment' (Gibbs and Simpson, 2004). It could be argued, however, that this is only true if learning is properly taken into account in the design of the assessment. Whilst there are many experiences of accidental learning or learning from unexpected inferences, these, by definition, are not the preferred model. If appropriate learning opportunities are not designed into the assessment then all it will actually do is *drive the students' activity* rather than their *learning*; a subtle yet important difference.

In addition to providing a 'learning opportunity', there are secondary issues connected to assessments, referred to as 'assessment backwash' (Biggs, 2003). This 'backwash' may, *inter alia*, provide students with a view of what the curriculum is; inform students as to how much time they are expected to spend on the module outside of the lecture theatre; and indicate which aspects of the module are the most, as well as which are the least, important components of the curriculum. All of this may be very different from the teaching team's perspective. This 'backwash' is an inescapable and integral feature of any

Table 1: Conditions of the Assessment Experience Questionnaire (AEQ)

Theme of the AEQ	Conditions of the AEQ
Time demands and student effort	Assessed tasks capture sufficient study time and effort. These tasks distribute student effort evenly across topics and weeks.
Assignments and learning	These tasks engage students in productive learning activity. Assessment communicates clear and high expectations to the students.
Quantity and timing of learning	Sufficient feedback is provided, both often enough and in enough detail. The feedback is provided quickly enough to be useful to students.
Quality of feedback	Feedback focuses on learning rather than on marks or students themselves. Feedback is linked to the purpose of the assignment and to criteria. Feedback is understandable to students, given their sophistication.
Use of feedback	Feedback is received by students and attended to. Feedback is acted upon by students to improve their work or learning.

Source: (Gibbs and Simpson, 2003)

assessment. When designing an assessment, therefore, there is a need to maximise the learning opportunities whilst also being aware of the backwash issues.

To help develop a credible assessment, recourse is also made to two evaluative instruments. First, a well-established guidance note setting out *Seven principles for good practice in undergraduate education* (Chickering and Gamson, 1987) and, second, an emerging evaluative instrument focusing solely on assessment – the Assessment Experience Questionnaire (AEQ) (Gibbs and Simpson, 2003). In the AEQ the researchers point to eleven assessment conditions clustered under five themes to support learning. For completeness, these eleven conditions, together with the Seven principles, are set out in Table 1 (see previous page) and Table 2 below.

As expected, there are many similarities in both instruments. Ensuring time-on-task, developing active learning techniques and setting high expectations are most notable in both, as is the value of prompt feedback.

An example of allowing students to act upon their feedback is given in Reynolds (2004). This approach adopts a so-called 'zero tolerance' approach to students' coursework submissions. In this example the students are required to re-submit their work, using the feedback, until the threshold pass criteria are met. This approach ensures that the feedback is used and acted upon and ultimately gives the students additional opportunities to demonstrate their new understanding.

In addition to this use of feedback to drive learning, there appears to be a resurgence in the use of formative assessment. In formative assessment the students' work is reviewed and

Table 2: Seven principles of good practice in undergraduate education

Principle	**Good practice…**
1	…encourages contact between students and faculty
2	…develops reciprocity and cooperation among students
3	…uses active learning techniques
4	…gives prompt feedback
5	…emphasises time-on-task
6	…communicates high expectations
7	…respects diverse talents and ways of learning

Source: (Chickering and Gamson, 1987)

commented on without a grade being provided; it is an opportunity for an activity to drive learning and not an activity for scoring marks. In this context students are free to expose their weaknesses and hence engage more openly in the learning process. With summative assessments, on the other hand, a final grade is awarded for the work. The summative approach suggests students are less likely to knowingly share their misunderstanding with their teachers, hence real learning opportunities may be missed (Biggs, 2003).

The general increase in formative assessment may be due to the increased opportunities provided by Computer Assisted Assessment (CAA). With mainstream and off-the-shelf CAA tools, teaching teams are able to create their own tests/quizzes using a range of different question types. Typically, these question types include multiple-choice, multi-select, hot spot or fill-in-the-blank questions. Tests constructed from such questions are generally referred to as objective tests and, in many cases, are generally taken online with marking and feedback often being instantaneous. Whole-class performance statistics can be produced to both to analyse students' learning and to provide a basis for the evaluation of teaching. A discussion of the design and use of objective tests can be found in Bull (2001), whereas examples of the benefits of such tests and their positive impact on student performance can be found in Smailes (2004).

Other benefits arise because most CAA is readily scalable, both up and down, and can, with careful forethought and planning, have a reasonable shelf life. After the initial pioneering work, CAA has now become a well-established area of educational enquiry. This is evidenced by the international conference on Computer Assisted Assessment which is now in its eighth year (Ashby, 2004).

There are concerns over the use of objective tests, however, and these generally revolve around answer guessing, answer sharing, the distracters (i.e. the alternatives in a multiple-choice question), giving inadvertent tips, and the biasing that may be associated with an online test (Brown et al, 1999). Moreover, some students may not be as computer literate as their peers and hence any online test may create an additional anxiety which may disadvantage them.

The modified assessment

To overcome the issues regarding poor student engagement and the failure of the phase test to encourage learning, a modified assessment regime was introduced in 2002/03. This modified assessment sought to exploit many of the Seven principles of good practice and the AEQ noted on the previous page, as well as having the fundamental purpose of improving the examination performance.

The modified assessment, discussed here, is a set of student-unique Weekly Assessed Tutorial Sheets (WATS) to replace the in-module phase test. The WATS are essentially a collection of weekly homework problems in the form of out-of-class tutorial sheets. Each WATS typically comprises a couple of key questions each of which has around four or five further questions. These sheets are marked and count towards the overall grade in the module. Although developed to support learning, this assessment is summative. Because the WATS are summative, a regular study pattern is 'enforced', which, it is hoped, will be seen as beneficial by the students and adopted in their other studies.

The student-unique nature of the WATS was achieved by embedding random factors into the questions. With this approach, the students could only discuss their methodology and not the answers. Benefits arise here because opportunities for student collusion and answer sharing are reduced, i.e. no two students will have the same answers. The students are still able to collaborate on methodology or contextualisation of the problem and this, it is hoped, will reinforce the idea and benefits of collaborative learning by allowing a group self-help culture. Furthermore, because each student tackles the same question, parity of assessment is assured. This is not the case in traditional objective tests which use question banks to deliver random questions to different students.

Naturally this regular and student-unique approach to assessment only becomes a viable proposition because it relies extensively on computers. Desktop software as well as bespoke software was used to write the tutorial sheets, deliver the work to the students, collect the students' answers and mark the students' work, thus providing prompt and individualised feedback and a group analysis of performance, together with a full worked solution. This fully integrated CAA programme was designed to promote learning and, of course, to counter the previous poor examination performance. What is attractive about this work is that it gains many of the noted benefits of CAA – speed of use, scalability, reuseability, rapidity of feedback, ease of student performance analysis – but it does not suffer from the noted problems of answer sharing or answer guessing and so on. Additional benefits also arise because the students download the WATS and tackle them in a setting and at a time that suits them. By definition, it does not tie them to a computer nor does it suffer from the previously mentioned disadvantages of computer anxiety and biasing.

One of the issues with many assessments, whether CAA or not, is that they often present one-off statements of learner performance and may do little to improve the learners' experience or their understanding of the subject. Such issues are particularly true for end-of-session summative assessments where opportunities for feedback or diagnosis of the students' learning have long since passed. Hence although CAA appears attractive, from a resource perspective, it does not necessarily imply a better educational experience. To counter such concerns these assessments (WATS) are continuous (weekly), and actively encourage the students to engage with the materials on a regular basis.

Many of the eleven AEQ conditions are apparent in the WATS assessment. The WATS capture *sufficient study time and effort*; by definition, they *distribute student effort evenly across topics and weeks; they engage students in productive learning activity*; and the WATS *communicate clear and high expectations to the students*. A further benefit of this approach is that it is consistent with the idea that a good curriculum has built-in opportunities for student practice (McAlpine, 2004).

One of the issues raised by the AEQ not presently covered by the WATS is the student response to, and use of, the feedback, i.e. no formalised mechanisms are currently in place to re-test the students' knowledge based on their individual feedback. This could be overcome by requiring the students to re-submit their work, after feedback, until a threshold performance level is met. An alternative approach could be delivering supportive remedial quizzes to the students via StudyNet. Although different, both approaches have the shared goal of closing the learning loop by re-evaluating the students' understanding post-feedback.

The Seven principles have *time on task, use of active learning techniques, setting high expectations,* and *prompt feedback* in common with the AEQ, which has already been discussed, but they also include the *importance of contact between students and staff* and *allowing reciprocity and cooperation among students* as well as *respecting diverse talents and ways of learning*. This assessment provides regular staff-student contact by delivering an individual task to the students and providing additional out-of-class advice via StudyNet. It also delivers a weekly and personalised email giving feedback on their performance. The WATS, whilst being a required task, has sufficient flexibility to respect the diversity of the ways of learning since, as discussed before, learners can tackle the work in the way that best suits them. For example, they can choose to work alone or with peers, they can work immediately or delay the activity until later in the week.

Impact of the modified assessment

Drawing guidance from the Seven principles and the AEQ, this work set out to develop a robust and learning-oriented assessment, and to improve understanding and ultimately examination performance. An overview of the impact of WATS on examination performance is given below in Table 3.

Table 3: Examination performance, pre- and post-WATS

	2001/02 (pre-WATS)	2002/03 (post-WATS)	2003/04 (post-WATS)
Mean	38.7	47.1	42.5
Median	34.0	48.0	44.0
Standard deviation	24.4	23.6	21.6
Number scoring over 35%	62	88	83
% scoring over 35%	49	67	65
Population	127	131	131

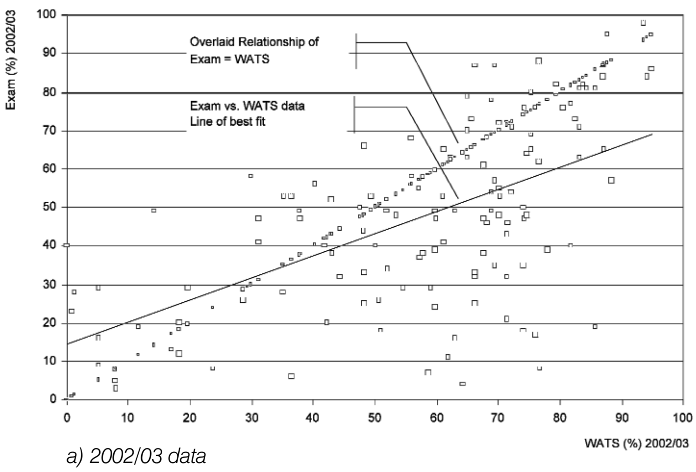
Table 3 shows the marked improvements in all the examination performance indicators. The mean examination score has improved from ~39% (2001/02 pre-WATS) to ~47% and ~43% in 2002/03 and 2003/04 respectively (post-WATS). The number of students scoring over 35% in the exam, the typical pass boundary, also rose from 62 (49%) (2001/02) to 88 (67%) and 83 (65%) after the WATS were introduced. This, it is hoped, will not only benefit this module but also the progression opportunities of the students and ultimately the student retention figures.

Having identified an improvement in examination performance, it is instructive now to see how the individual students' WATS scores compared with their exam scores, the thesis being that a high WATS score would also be followed by a high exam score. A plot of the individual students' exam scores versus their WATS scores is shown in Figure 1, for both 2002/03 and 2003/04. By setting and drawing a relationship of exam score = WATS score, as denoted by squares, it becomes easy to see which assessment appears to be most favoured by the students, i.e. which students scored better in the exam (circles above this exam score = WATS score relationship) and which students did better in the WATS (circles below the exam score = WATS score relationship). In both instances 2002/03 and 2003/04 most students performed better in the WATS than they did in the exam, i.e. 88/131 (~67%) in 2002/03 and 116/131 (~88%) in 2003/04.

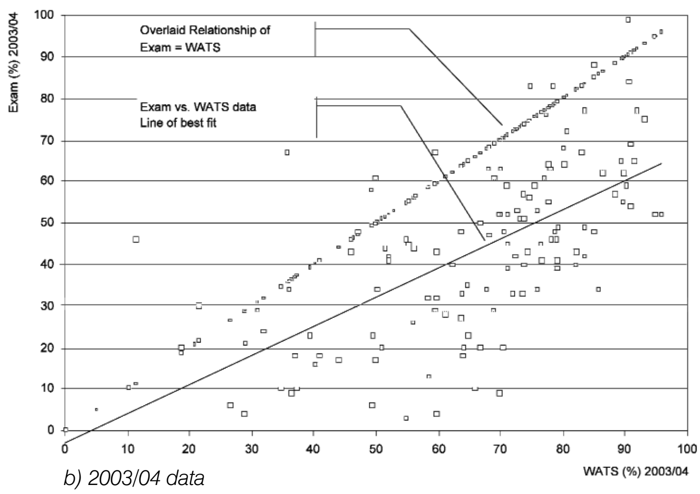
This may be due to the less stressful nature of the WATS – an in-module assessment. Here, for instance, they can talk with their peers, seek tutor guidance and use their module notes as well as other support materials to help them. The examination, on the other hand, is a traditional, closed book, formal assessment; such tests often return lower grades than the more open, in-module assessments – this thesis is also supported by Snyder (1971). It is also possible that the WATS were easier than the examination questions and whilst there are no major concerns about this – the motivation for the WATS was to provide practice opportunity and develop the students' confidence and understanding in the subject – making the WATS too easy may have inadvertently fostered a false sense of security, another feature of assessment backwash.

Although difficult to prove a trend from the two sets of data, i.e. 2002/03 and 2003/04, more students did better in the WATS than in the exam in 2003/04 than they did in the previous year.

Figure 1: Scatter plot of individual students' exam and corresponding WATS performance



a) 2002/03 data



b) 2003/04 data

A measure of this cohort difference between WATS and exam can be made by calculating the Sum Absolute Difference (SAD) and the Average Absolute Difference (AAD) – see Table 4 below. For the sake of completeness, details of the AAD and SAD calculations are provided in an endnote.

Table 4: Calculations of the Sum and Average Absolute WATS – exam differences

	2002/03	2003/04
Sum Absolute Difference	2,545	3,661
Average Absolute Difference	19	25

In both tests, AAD and SAD, a low score indicates a better correlation between the WATS and the exam. Hence not only did more students in 2002/03 do better in the exam than in the WATS compared with those in 2003/04, but the 2002/03 class as a whole, as indicated by the AAD and SAD scores, have a better overall correlation between WATS and exam scores. This suggests that this approach was more effective in its first year. One reason for this may be due to the fact that the 2002/03 WATS questions were re-used in 2003/04. The major effort in 2003/04 was the development of the automation facilities and not rewriting the WATS questions. A student from 2002/03 could have passed the full WATS solutions (uploaded to StudyNet after the submission deadline had passed) to a 2003/04 student and they could then have found their way to others in the cohort. Sharing the solutions would not have helped the learning process and may also have developed in these students a false sense of security. To overcome this possibility a new WATS question set will be developed for 2004/05.

Discussion and conclusion

It is apparent that assessment has a major impact on student activity and ultimately learning. This work has shown that by replacing a 'one-off' form of assessment (i.e. an in-module phase test) with a series of weekly-assessed tutorial sheets, a significant and positive impact on the students' examination performance and hence understanding of the subject can be gained.

The reasons for these improvements are believed to be because of the close alignment of this work with the AEQ and the Seven principles previously mentioned. The work set out to develop learners and not students simply seeking to gain marks. Some of their on going learning would have undoubtedly helped them in their understanding of the subject and ultimately their examination performance.

There are still many questions resulting from this work and these will be explored in further investigations. Outstanding questions include how a student can score >85% in the WATS and yet still not pass the exam, and what strategies a student was adopting to score only ~15% in the WATS and yet manage to get ~50% in the examination.

Although steeped in an engineering context, there are many general and transferable themes that emerge from this work. The benefits of regular assessment, the delivery of student-unique work, the use of computers to provide personal and rapid feedback and the organisation of the assessment around good educational practices are all worth considering in any discipline.

Endnote

[1] Equations for the Sum Absolute Difference (SAD) and Absolute Difference (AAD):

$$SAD = \sum_{n=1}^{Ns} \left| WATS - Exam \right| \quad \text{Eq 1}$$

$$AAD = \frac{\sum_{n=1}^{Ns} \left| WATS - Exam \right|}{Ns} \quad \text{Eq 2}$$

Ns = the number of students in the cohort, which for both cases Ns = 131

References

Ashby, M. (2004) *Proceedings of the 8th Computer Assisted Assessment Conference.* Loughborough

Biggs, J. (2003) *Teaching for Quality Learning at University.* Buckingham: Society for Research in Higher Education and Open University Press

Brown, S., Race, P. and Bull, J.M. (1999) 'Computer anxiety in students: Should computer-based assessment be used at all?' in *Computer-assisted assessment in higher education.* SEDA

Bull, J. (2001) *Designing and Using Objective tests.* Computer-Assisted Assessment Centre

Chickering, A.W. and Gamson, Z.F. (1987) 'Seven principles for good practice in undergraduate education' in *AAHE Bulletin,* Vol 39, Issue 7, 3-7

Gibbs, G. and Simpson, C. (2003) 'Measuring the response of students to assessment: the Assessment Experience Questionnaire'. 11th Improving Student Learning Symposium, Leicestershire

Gibbs, G. and Simpson, C. (2004) 'Does your assessment support your students' learning?' http://www.open.ac.uk/science/fdtl/documents/lit-review.pdf (accessed 4 March 2005)

Green, D.R., Harrison, M.C. and Ward, J.P., (2003) Mathematics for Engineers – The HELM Project. Proceedings of the Progress 3 Conference: Strategies for student achievement in engineering. Hull

McAlpine, L. (2004) 'Designing learning as well as teaching' in Active Learning in Higher Education, Vol 5, Issue 2. London: Sage Publications

Ramsden, P. (1992) *Learning to teach in higher education*, pp.187-88. London: Routledge

Reynolds, J. (2004) *Evaluation of a zero tolerance assessment strategy for incorporating health and safety into an undergradute Civil Engineering course.* Learning and Teaching Support Network (LTSN) Engineering subject centre. Mini-project report

Russell, M.B. (2003) *Using discussion forums to support teaching and learning.* Proceedings of the Progress 3 Conference: Strategies for student achievement in engineering. Hull

Snyder, B.R. (1971) *The hidden curriculum.* Cambridge MA: MIT Press

Smailes, J. (2004) 'Forcing the learning: observations from the implementation of a regular, computer assisted assessment strategy' in *Researching the challenges in academic practice.* University of Northumbria Press

Williamson, S. and Green, D. (2004) *Learning and Teaching Support Network: Engineering mathematics examples and case studies swap-shop.* Loughborough

Biographical notes

Mark Russell has been teaching at the University for nine years. His interests lie in engineering education and using a range of educational settings to both support and challenge his students. Mark was awarded the *Times Higher Education Supplement*/LTSN Generic Centre E-tutor of the Year in 2003 and more recently was one of the first five recipients of the Vice Chancellor's Award for Excellence in Learning and Teaching 2003/04. He has published papers in his subject area, computational fluid dynamics, and has also presented at various national and international conferences in computer-assisted assessment and engineering education.