Closing the Gap between Assessment and Learning

Sarah Maughan
University of Cambridge International Examinations
UK

Abstract
High stakes assessment has always played an important role in education, originally primarily for selection purposes and monitoring school performance over time. Whilst these two purposes continue to be important, increasingly we see added to these the aims of raising standards by improving the effectiveness of individual learning, and of encouraging the development of higher order skills such as Problem Solving and Creativity. The new aims of assessment show clearly how it can now be designed to impact on the actual learning that takes place in schools rather than purely to measure the outcomes of that learning.

This paper will present the ways in which the University of Cambridge International Examinations (CIE) is responding to this change in emphasis. It will give detailed examples of some assessment developments that aim to have a positive backwash on the teaching and learning that takes place in schools, including diagnostic testing, school-based assessment components and authentic assessments. The ways in which new technology can be used to assess process, as well as the traditional assessment of outcome, will be discussed.

The methods by which the boundaries between assessment and classroom practice are becoming blurred will be highlighted through each example.

1 Background

Public examinations started in the UK in the 19th century, as a means of selecting a small group from an elite upper class to go on to further study. Over one hundred years later the primary aim of examinations was essentially unchanged; however, a number of new uses were beginning to emerge. Government policies were directed towards education to increase access to a wider proportion of the population, and to encourage economic growth by providing a highly educated work force. These policies led to a greater need for a reliable and valid system of assessment: 'comparative study of examination systems find its strongest justification when nations seek ways to bring about educational change, especially when the concern focuses on raising the level of school achievement' (Eckstein and Noah, 1993).

Increased numbers of students were encouraged to stay on in school to receive a more in-depth education, and examinations were used to measure whether performance in schools was improving with the different educational reforms and initiatives. The changing environment saw the desire for new skills to be taught to the student population including problem solving, creative thinking and the flexibility to learn and re-learn. Black et al (2003) state that developing students as 'successful learners' 'is a feature of great importance, for any exploration of the ways in which society is likely to change in the future is bound to point to the overwhelming importance of the capability to learn anew, and then re-learn, as rapid changes overtake us'.

The high stakes examination system in the 1980s and 1990s became seen as a catalyst that would force the necessary changes in the schools. 'Examination systems strongly influence the school curriculum, to the extent that in England/ Wales, France, China and Japan the subjects examined and the syllabi for the examinations virtually determine the school curriculum and the objectives of teaching. The format and content of examinations reflect what an educational system considers to be the knowledge and skills of most worth.' (Noah and Eckstein, 1993)

As Raffan said at the first ACEAB conference in 2000, 'there is no doubt that the techniques and frequency of assessment and examinations do profoundly affect both the content of the curriculum and how it is taught'. So how have the nature and content of examinations changed

over the years to meet the changing uses to which they are put, and to encourage changes to what goes on in schools?

This paper describes the recent history of syllabus development in two subjects from the University of Cambridge International Examinations[1]: Biology and Geography for students in the 14 – 19 age group, and through them highlights how the nature of examinations has and is still changing to meet requirements.

The relatively recent introduction of the use of new technologies in schools and in examinations is providing the opportunity to make additional and more rapid changes and some of the impacts are discussed in Section 3 of the paper. The changes that new technologies enable are part of an ongoing process that could see examinations change dramatically in the relatively near future from the way they have been for over 100 years.

In our paper 'Assessing 'New Educational Goals' presented at the IAEA conference in Manchester (Puntis et al, 2003) we defined a new paradigm for assessment in the 21st century and provided case studies to exemplify its parts. The new paradigm is based on five key features:

1 authenticity,
2 systemic validity,
3 focus on 'deep understanding',
4 focus on skills,
5 duality of purpose.

This paper will develop this theme to show how assessments that have these features are part of this process of change in examinations, and impact in a positive way on learning and the curriculum.

2 Diagnostic Testing to Modify the Learning Process

'Assessment can give insights into very specific aspects of the thinking and performance of pupils. How are they thinking about a particular situation? Where and why is skilled performance on some task breaking down? Using assessment to ask and answer such questions improves the information available to the teacher and makes it possible to identify and address learning difficulties.' (Raffan, 2000). This new purpose of assessment, to produce feedback for the teacher to inform the learning, is one of the areas in which there have been significant developments and in which further progress is likely to be made. An example of one such development from CIE is detailed below.

2.1 Cambridge Checkpoint

Cambridge Checkpoint was introduced in May 1999 and was a new departure for CIE in that it aims to assess students **prior** to their commencing a course leading to a recognised qualification. Results highlight students' strengths and weaknesses to allow the future learning to be modified appropriately. The tests were designed against a tight specification which ensured that different content and skill areas were targeted in such a way that meaningful feedback could be produced against each area. Items were pre-calibrated to ensure that the tests targeted the appropriate ability ranges and to allow feedback to be given much more rapidly than in the usual examination processes. A curriculum framework, rather than a complete syllabus, was produced against which the tests were constructed. This provided schools with the opportunity to use the tests after following a range of different curricula in the 11 to 14 stage. The tests targeted the core subjects of English, Mathematics and Science. This paper focuses on the development of Biology assessment in the recent past so the following examples have been taken from Checkpoint Science.

[1] NB The University of Cambridge International Examination (CIE) was established in 1998, to be responsible for the international examinations of the University of Cambridge Local Examinations Syndicate (UCLES).

On completion of the tests the students are given a 'broad brush' report on their performance in the main areas of the subject. The tables below show the results for an individual student.

Table 1: Individual Student Feedback from the Checkpoint Tests

Overall Result		Topic Results	
Subject: Science		Biology:	2.6
Checkpoint Score: 2.1		Chemistry:	1.5
Date: September 2003		Physics:	2

What you got right and what you got wrong

Most of your answers were as expected for a student with your Checkpoint score. However, some of your answers were surprising: the most surprising ones are listed below. They may give you information about what parts of the subject you are good at and what parts you need to w ork harder at.

Questions you answered correctly that we expected you to find difficult

Question and part	Topic	Sub-topic
5b	Physics	Light
1bii	Physics	Electricity
1b	Biology	Humans as organisms
7cii	Chemistry	States of matter and physical change

Questions you answered incorrectly that we expected you to find easy

Question and part	Topic	Sub-topic
6a	Chemistry	Materials
6a	Biology	Variation and classification
11c	Biology	Ecosystems
8b	Physics	Measurement and properties of matter

This feedback, although limited, begins to form the basis of a dialogue between the student and teacher about future learning routes and areas to concentrate on. In this way the tests could be used to provide formative information and as a basis on which to develop individualised learning plans for students. The results could also be used to choose the most appropriate course for students in the run up to the more formal, often tiered examinations at 16.

In addition to the reports for individual students are reports for different class groups, and for whole schools, and each of these are mapped against the benchmark of the whole cohort performance. The table below shows a sample taken from a class report.

Table 2: Class Feedback from the Checkpoint Tests

Average Checkpoint scores are as follows:

Science (overall)	Average Checkpoint score = 3.2
Biology	Average Checkpoint score = 3.1
Chemistry	Average Checkpoint score = 2.7
Physics	Average Checkpoint score = 3.6

Explanation of the Checkpoint Scale:

- Scores on the Checkpoint scale are from 0.0 (the lowest level of ability) to 6.0 (the highest level of ability).
- The standard of performance represented by a number on the Checkpoint scale is the same in every session.
- An "average" Checkpoint student should achieve a score between 3.0 and 4.0.
- The proportion of students achieving scores between 2.0 and 5.0 is usually about 75%.
- The scale is the same for the subject as a whole and for each topic.

Note:

Any student who missed a paper or who obtained a score of zero on either paper has been omitted from the analysis on this page.

Performance in each sub-topic

For this part of the report, we have used the overall ability of each student and the difficulty of each question to predict how well each student should perform on each question. We have then compared their predicted performance with their actual performance. If there is a sub-topic where your group of students has performed better than expected, this may indicate that the teaching has been particularly effective in that sub-topic. If there is a sub-topic where your group of students has performed worse than expected, this ***may*** indicate that the coverage of the sub-topic was incomplete or that learning in this area has been less thorough than in other areas.

Biology		
Bc	Cells and Organisms	Performance was as expected
Bh	Humans as Organisms	Performance was as expected
Bp	Plants	Performance was worse than expected
Bv	Variation and Classification	Performance was as expected
Be	Ecosystems	Performance was as expected
Chemistry		
Cm	Materials	Performance was worse than expected
Cs	States of Matter and Physical Change	Performance was as expected
Cc	Chemical Change	Performance was as expected
Cp	Periodic Table	Performance was as expected
Physics		
Pp	Measurement and Properties of Matter	Performance was as expected
Pf	Forces and Motion	Performance was as expected
Pe	Energy	Performance was as expected
Pl	Light	Performance was as expected
Ps	Sound	Performance was better than expected
Pm	Magnetism	Performance was as expected
Pc	Electricity	Performance was as expected

At this level the feedback is more conclusive than that provided for individual students. Statements describe strengths and weaknesses rather than merely stating which questions were right and wrong. This becomes possible because the results are more reliable when they are from a number of students for each item rather than the results of only one student. The teacher can use this feedback to tailor the future course to the particular needs of the group.

2.2 Discussion

Tests of this type highlight one of the ways in which the new purposes of assessment are being realised by examination boards. Tests are being used to inform aspects of the learning process; and this clearly shows one way in which the boundary between learning and testing is becoming blurred. The Checkpoint tests make explicit the aim to influence the learning that takes place in schools, and expects the teaching to be modified in light of the feedback. Although in this case the nature of the questions themselves remain similar to previous examinations, the purposes and the amount of information produced as a result have changed.

It is the impact on the school curriculum that appears to have been of most use to many users rather than individualised learning plans. A number of schools do not use the tests on a regular basis, with every year group, as would be expected if individual learning was seen as key. The tests are often used as a less regular measure, to compare a school's performance with that of the international cohort, with the view to adapting their curriculum, and re-using the tests at a future point to measure the effectiveness of any changes. This use of the tests as a monitoring tool, to benchmark performance against an international cohort and then to measure effects of policy change, either within a school or within a country, has been found very useful by a number of users.

For assessment to have a greater impact on the curriculum and on learning the provision of a single mark or grade is insufficient information. When the primary aim of assessment was to select students to go on to the next stage then this means of providing results suited the purpose. The new uses of assessment results, such as to encourage a positive backwash on the learning, require more detailed feedback. As mentioned before, CIE introduced its first diagnostic test in 1999, which can be seen as a key departure for an examination board, and a recognition of the change in the function of examinations, and a recognition that we must change what we do in order to respond to that.

3 School-based Assessment in the University of Cambridge International Examinations Syllabuses

One major theme of the ACEAB conference in 2000 was school-based assessment. This form of assessment has been a relatively recent addition to examination systems in many countries and the many positives, as well as the difficulties, associated with its introduction are an interesting source of discussion among examination boards. The following section briefly describes how the use of coursework components has developed in UCLES examinations in general, and CIE examinations in particular, over the recent past.

In 1988 a new suite of examinations; GCSEs (General Certificate of Secondary Education), was introduced into the UK education system as a replacement for the previous examinations for 16 year old students. At the same time UCLES introduced its International GCSE (IGCSE) to meet the needs of some of its overseas users[2]. The previous system had been made up of two sets of examinations, CSE for lower ability students, and O Level for higher ability students. The introduction of the new examinations, both GCSE and IGCSE, addressed a number of needs. While some commentators have noted that the GCSE introduced a centralised control over the school curriculum (e.g. Eckstein and Noah, 1993), more significantly the development enabled a more inclusive form of assessment in schools, positively rewarding the achievements of all students, and a greater focus on the assessment (and therefore teaching) of skills in schools. In the CIE Science syllabuses this meant a greater emphasis on the questions assessing the application of knowledge in written papers, and a greater emphasis on the development of practical skills throughout the syllabuses and in the curriculum.

3.1 School-based Assessment in Biology

Prior to the introduction of the GCSE and IGCSE in Biology there had been a limited mention of skills development within the syllabus and a similarly limited attention to the teaching and learning of these skills in schools.

The assessment objectives in the syllabus now have a much clearer emphasis on the development of skills, and further weight is given to this in the supplementary text such as 'These skills cannot be precisely specified in the curriculum content as questions testing such skills are often based on information which is unfamiliar to the candidate. In answering such questions, candidates are required to use principles and concepts that are within the syllabus and apply them in a logical, deductive manner to a novel situation.' (taken from the 2005 IGCSE Biology syllabus)

At the time of the introduction of GCSEs, coursework components were added, as a compulsory component in the UK, and as an optional component in the international syllabuses. This optionality in the IGCSE was in response to customer needs, in some Centres it was difficult to access materials and resources that were needed to undertake a comprehensive program of practical work, and for some larger customers, compulsory practical work during a course was not a requirement. However, the emphasis on these parts of the syllabus remained, whatever the means of assessment. It was important for us to be confident that each student experiences a 'consistent' experience and is assessed on the same assessment objectives[3].

The 2005 IGCSE Biology syllabus states 'Scientific subjects are, by their nature, experimental. It is, accordingly, important that an assessment of a student's knowledge and understanding of

[2] NB UCLES also continued to offer the O Level examination for customers who wanted an assessment specifically for the higher ability range. O Levels have also undergone extensive syllabus change since 1988 and now include a greater proportion of items assessing application of skills. However, O Levels continue to be assessed using paper-based examinations at the end of a course, school-based assessment components are not generally available.

[3] Practical skills are assessed via a school-based component or via an end of course practical examination, and in a small number of cases, where other methods are not possible, as an alternative to coursework written paper.

Biology should contain a component relating to practical work and experimental skills'. In this way, the emphasis on practical work in the syllabus aims to have a positive backwash on the learning that takes place in schools, and have a greater impact on the learning than was apparent in the earlier syllabuses.

Teachers play a key role in the school-based assessment component of the examination[4] and as such they must be accredited by Cambridge before commencing teaching of the course. Accreditation can be gained by past experience, but in most cases it is achieved by attending a face-to-face training course or by completing the Coursework Training Handbook. This Handbook recommends appropriate tasks as guidance to the teacher but does not define the tasks that must be undertaken, this is left flexible for the teacher to decide what is appropriate for the group and in their context. The Handbook also describes the ways in which feedback can be given to the students as part of the assessment process, to scaffold the learning that is taking place without providing the answers to the assessment itself, before going on to describe the standards that must be achieved. Coursework sat in this way makes up 20% of the final assessment grade. The training that is provided to the teacher is a key element in the implementation of school-based assessment. The establishment of teachers as 'assessment partners' aims again to have a positive influence on the learning process, enabling assessments to be personalised to the group of students, and again providing a basis for dialogue between the teacher and student on the learning that is taking place.

3.2 School-based Assessment to Develop Geographical Skills

As with practical work in Science, field work has been a key component of learning Geography for many years but an explicit part of Geographical syllabuses only since the introduction of GCSE in 1988[5]. In the GCE O Level Geography syllabus for 1985 the references to skills were limited to the very start of the document which stated 'Candidates should, where possible, make a study of their own home area or of some area of which they can have first-hand knowledge: additional credit will be given for the application of this knowledge in the examination'. A project was available as an optional component of the examination, similar in length to the combination of the two pieces of work that make up today's coursework option. However, directions as to the aims or assessment objectives of the option were limited, merely adding 'Projects … should include some evidence of personal observation and recording of data and some analysis of the collected information. Projects should either attempt to answer a question or discuss a problem'.

The introduction of the IGCSE in Geography marked an increased emphasis on the assessment of Geographical skills of data collection, measurement and presentation of results. Assessment objectives in the 2005 syllabus are:

A Knowledge with understanding
B Analysis
C Judgement and decision making
D Investigation (enquiry skills, practical skills and presentation skills)

More recently this skills development has progressed further, in some syllabuses, into the areas of hypothesis testing and decision making. As with the IGCSE in Biology, it did not prove possible for CIE, with its diverse international market, to introduce compulsory coursework to its Geography syllabus, although the syllabus does place an emphasis on the associated skills[6].

[4] For the school-based assessment, practical tasks are set by the teacher throughout the course and are assessed by the teacher. A sample of coursework is supplied to Cambridge for moderation, for each student the work must meet a tight specification, demonstrating ability in a number of different content and skill areas of the curriculum.

[5] A limited number of earlier Geography syllabuses required the submission of a field work folder, though this tended to be descriptive rather than analytical. There was also some limited use of 'skills-based' papers, incorporating such skills as atlas work.

[6] A paper-based alternative to coursework is available at the end of the course, alongside the other written papers. The paper has the same balance of assessment objectives as the coursework component.

The emphasis on these skills, whatever the means of assessment, is intended to change the emphasis within the schools, and have an effect on the curriculum and the learning process.

The emphasis on skills development is also apparent in the non-coursework papers, and the description of the current paper 2 is that it 'will be entirely skills-based and will test a candidate's ability to handle various ways of depicting geographical information … Candidates will be able to demonstrate skills of analysis and interpretation and application of graphical and other techniques as appropriate'.

Unlike the Science syllabus in which the coursework components are made up of a number of pieces of work, each targeting particular skills or syllabus content, the Geography coursework component is assessed through two longer pieces of work through which all skills must be demonstrated by the student. To enable students to approach tasks appropriate to them in terms of context and culture, the titles are selected by the teacher, with approval of the Moderator. In this limited way the work can be individualised to different students' interests and environments, and localised to the requirements of particular countries or cultures.

3.3 Discussion

School-based assessments are used in many countries and a variety of arguments are given in their favour, a common theme to all arguments is the impact that the use of such assessments will have on the curriculum, and the teaching and learning. It therefore follows that the nature of assessment in many countries is changing to meet the changing uses to which results are put and assessment is both affected by the uses and used as a driver for change. The reasons given at the first ACEAB conference for introducing coursework are fairly representative of those given in other contexts and can be grouped into four areas.

1 The first reason for the introduction of coursework is the ability to include a wider range of skills in the high stakes assessment. As Berry (2003) pointed out, 'Using varied assessment tasks, for example, portfolios, observations, experiments, projects, simulations, interviews, performances, presentations, peer-assessment, and self-assessment allows a deeper understanding of students' learning in different perspectives' and clearly using school-based assessment does add variety to the assessment process. The following quote from the Namibian perspective is a very clear example of this. 'CA was envisaged to broaden the scope of assessment to include progressive or alternative forms of assessment that will create a wider spectrum of opportunities and contexts for learners to demonstrate what they know and can do. In this context CA was seen to promote learner centred education.' (vd Merwe, 2000). A similar perspective was given from Malta, 'From an educational point of view, the justification of the school-based component is that it makes the final assessment more valid because of the range of skills and processes that could be included, and the range of contexts in which the assessment could take place.' (Grima and Ventura, 2000)

2 A second common reason given for developing a school-based assessment system is the professionalisation of teachers by including them more in the testing process. Again, Namibia places an emphasis on this when explaining the reasons behind introducing such a system. To change the very high drop out rate, and teaching and learning practices resulting in rote learning and memorization 'it was felt that curriculum-based CA could make a difference to improve the quality of teaching and assessment practice.' (vd Merwe, 2000)

3 The third reason in favour of the introduction of school-based assessment is the motivational factor that the personalisation of the learning will afford. When teachers and students can select contexts or titles that reflect their own interests or environments then it is hoped that it is more engaging: 'In effect, school-based assessment becomes an engine for enriching the curriculum as it is delivered. Not only is there better correspondence between what is done in schools and what is examined, which itself enhances validity, but the result is fuller and richer.' (Wood, 1991)

4 Finally the use of school-based assessment is seen as a means of developing students' ability to learn and to reflect on their own learning; in order to develop flexible members of

> a rapidly changing world of employment. Again, Namibia see this as a key reason for introducing school-based assessment systems; believing they 'should help teachers to put far greater emphasis on learning which can be transformed and applied to new circumstances than on learning facts and procedures applicable only in situations closely similar to those in which they were learnt.' (vd Merwe, 2000)

Despite the many benefits of school-based assessment there remain a number of difficulties associated with its introduction, including difficulties with training teachers, which it is commonly accepted is a large scale and long-term issue, workload for both teachers and students, and ensuring the integrity of the work. An article in the Times Educational Supplement quoted one Head of Department from a UK school as saying he would willingly move away from coursework because 'it distorts learning, adds needlessly to students' and teachers' workloads and is open to abuse'. (Mansell, 2003)

A further debate is underway in the UK over the usefulness of students undertaking school-based assessment in a number of subjects, when essentially the skills that are assessed in the many pieces of work may be the same. Mike Tomlinson, who has been tasked by the UK Government to undertake a review of 14 – 19 education, was quoted as saying it is nonsense to have someone doing eight separate pieces of assessed coursework for their GCSEs, when just two pieces, testing across the curriculum, might be more appropriate, (Mansell, 2003). In the section below it is discussed how skills within a course can be assessed successfully through computer-based assessment, alleviating some of the difficulties of school-based assessment while still maintaining many of the benefits.

4 The use of New Technology to Assess Process and to Improve Authenticity

In the introduction the use of new technology was described as a means of speeding up and further changing the nature of examinations. The paper as a whole argues that examinations have changed in purpose and now have a greater impact on the curriculum and learning and this section describes how technology facilitates this. Both the Biology practical coursework assessment and the Geography field work coursework component for IGCSE have a paper-based alternative, for those users who are unable to undertake school-based assessment. The positive benefits of school-based assessment in terms of the impact on the curriculum and the emphasis on skills development are clear and these impacts are likely to be limited with the paper-based alternative. It is in this area that new technologies may enable us to gain similar benefits without offering coursework options. In fact in some areas it may offer benefits over the school-based assessment or the practical examinations.

In CIE A Level Biology practical assessments[7] a variety of performance tasks are presented to students targeting the assessment objectives of handling, measurement, observation and interpretive activities. One of the tasks is usually based on microscopy, as a context for assessing all these objectives. CIE is investigating the use of new technology to increase the range of skills that can be assessed in final examinations, without the difficulties associated with coursework components or practical exams. This does not mean the developments will replace practical work in either form, and in many cases this may remain the most appropriate form of assessment, neither does it mean that we want the use of practical work in schools to decrease. We hope that this approach to the end of course assessment will lead to an increase of practical work or field work during the course.

New technology provides the opportunity to simulate real tasks in a controlled computer-based environment. It is hoped that the simulations developed enable skills and constructs to be assessed in an environment as close as possible to the real thing, thereby providing a valid and authentic assessment of the targeted objectives. Kane et al (1999) describe authentic assessment as involving a 'close similarity between the type of performance which is actually

[7] Coursework options are not available in CIE AS/ A Level examinations so the skills are targeted in compulsory practical components.

observed and the type of performance which is of interest' and this is exactly what we hope to achieve by using simulations.

4.1 The Bioscope

One development of a subject-based simulation that CIE is piloting is the use of a computer-based microscope as part of an alternative to practical examination. This simulation has been developed with the University of Derby, Centre for Interactive Assessment Development (CIAD) and mimics the use of a 'real' microscope in a computer-based environment. The simulation is not intended to replace the use of a microscope during the course of study but provides an environment in which the skills needed to use a microscope, such as magnification and focusing, can be developed. The simulation enables groups of students to view and discuss the slides and work collaboratively to respond to the tasks as part of the learning process – something that is less easy to do with real microscopes. It is anticipated that this will be used to supplement work carried out using a real microscope.

The use of the microscope simulation has a number of advantages over its real alternative in an assessment environment allowing the stated assessment objectives to be assessed more directly and improving the reliability of the assessment. The simulation ensures that each student has the same experience – the slide will be identical for all users and on each occasion, something that cannot be guaranteed with the use of real slides and microscopes. However, perhaps the most important way it improves the assessment is by allowing us to track how the student has arrived at their answer. In current paper-based examinations it is not possible to ensure that students have really identified a particular structure in order to draw it, or magnified it or focused it appropriately, or even to be sure that the student has not merely made a drawing from memory of similar drawings in books. These things become possible with the simulation, ensuring that student performance is assessed rather than rote learning. The diagram below shows a question in which the students must click on the target structure before they draw it and this snap shot is used as part of the assessment.

Image 1: Bioscope with integrated computer-based assessment.

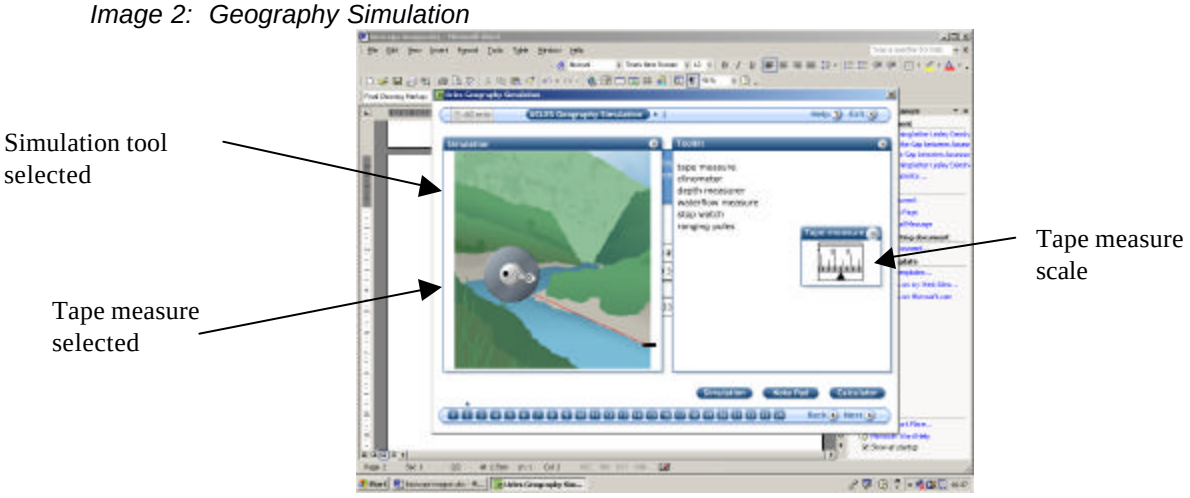
4.2 The use of New Technology to Assess Process and to Improve Authenticity in Geography

As in the Biology syllabus we described above, the alternative to coursework written papers in Geography are not seen as the ideal way in which to impact on learning or to assess the required skills. For this reason we are also investigating the use of new technology to improve the validity

and authenticity of the alternative to coursework paper in Geography. There are currently three simulations in development, in partnership with a UK-based technology provider, BTL. The three simulations cover the most popular areas which students focus on in field work: 'rivers', 'weather' and 'inner cities'. This range is particularly useful for CIE's diverse user base, some of whom may find it difficult to visit 'inner cities' so any learning in this context is, therefore, totally theoretical, similarly some Centres are based a long way from rivers or coastal areas so practical work focusing on these types of environment is not possible.

The simulations enable us to set tasks which are practical in nature and which assess skills such as hypothesis testing and development, decision making and so on. The current tasks have been written as hour-long assessments each developing certain ideas and themes linked around the same context. The simulations enable students to select tools, use them to make the necessary measurements in the ways which they think best, and use their findings to produce graphs and draw conclusions.

Image 2: Geography Simulation



This type of activity has been imitated in written papers in the past but the fact that the students must now actually 'do it' as part of their examination is likely to provide significant improvements in the validity of assessment. Pursuing this idea further it may be argued that the use of the simulations as part of high stakes assessment has some advantages over the use of field trips as a coursework option. Using the latter methods the assessment is based on a written report that the student produces, on their practical experience, rather than the practical work itself, therefore the validity of the assessment is limited. Field work is still seen as an essential part of a good Geography course, but it does not necessarily follow that the best form of assessment is a written report on that work.

4.3 Discussion

The development of simulated environments for use in assessment is also an area where the learning and assessment boundaries become unclear. 'The incorporation of technology into assessment is inevitable because, as technology becomes intertwined with what and how students learn, the means we use to document achievement must keep pace.' (Bennett, 2002) It is known that students who sit paper-based examinations after a course that integrates new technologies, tend to under-perform in the final exam (Russell and Plati, 2001). Similarly it could be expected that students who use computers for assessment will under-perform if they have not used computers during the course. If we are to use a computer-based simulation as part of assessment then it must be developed for use during the learning also – and this has in fact been done with the computer-based microscope. The Bioscope has been designed to operate in both assessment and learning modes. In assessment mode a combination of computer- and paper-based tasks are presented to the student around a set of slides, as in a practical examination.

The objective computer-based tasks are marked by the computer, enabling immediate feedback at a detailed level, to be provided to students, and relieving teachers of some of the more administrative aspects of marking. This version of the Bioscope is being piloted with students in April 2004 to investigate its potential for use in high stakes exams. As with diagnostic tests and school-based assessment, the feedback provided from such assessment used in a low stakes environment could be used formatively with the students. This information could then be used to modify the learning.

The learning version of the Bioscope contains slides that cover the complete range required for an A Level-type course, both in terms of content and the skills and concepts that are to be developed. There is a total of 40 hours worth of tasks on the CD, and these tasks are provided in Word format as well as pdf, so that teachers can adapt them to their own and their students' requirements. It is also expected that teachers will use the slides for their own tasks and to develop skills that they see as required by their students. This learning version of the microscope will shortly be published by CUP, with learning tasks that cover all the requirements of the CIE A Level Biology syllabus and, where appropriate, IGCSE and O Level syllabuses.

It is hoped that there is an impact on the learning in both modes, feedback from assessment used as a low stakes tool, and development of skills in a structured practical environment to allow greater exposure and to scaffold learning. Also it is anticipated that the use of the simulations could enhance collaborative working and the skills associated with this.

Like the microscope simulation, the Geography simulations as part of an assessment are being piloted with students in 2004, however, also like the microscope it is hoped that ultimately the simulations will be used by students to develop their learning. For practical reasons, if the student is to use a simulation as part of an assessment it is important that they have used it, or similar simulations, previously, to ensure they are familiar with the interface, however, the simulations will also be useful tools for teachers. It is hoped that they can use these to develop Geographical skills in students in an interactive and engaging way. It is also possible to use the simulations to develop collaborative working, in learning mode.

By developing the simulations CIE is again aiming to have a positive impact on the curriculum; if the assessment involves 'doing' and applying knowledge in an interactive environment, it is hoped that the curriculum will be adapted to include more learning of these skills, thereby developing more flexible, reflective learners.

5 Conclusions

This paper aims to demonstrate some ways in which examinations have changed over the recent past; explicitly aiming to have a greater impact on the curriculum and on learning. Examples are given of syllabus developments which characterise these changes; including the development of diagnostic assessments, the use of school-based assessment and the use of computer-based simulations. A new assessment paradigm was suggested by Puntis, Beedle and Maughan in 2003, involving a focus on authenticity, systemic validity, deep understanding, skills development and duality of purpose. It is argued that the changes in assessments that follow on from an acceptance of this new paradigm lead to a blurring of the boundaries between assessment and learning.

The paper describes some ways in which assessments with these features affect learning and the curriculum, they encourage 'doing' and the measurement of 'doing' rather than memorisation, they encourage reflection by students on their own learning, and they development flexible learners who can transfer skills and knowledge to new contexts. New technology, in particular, offers new opportunities – not for all subjects or in all contexts, but where it can offer advantages over existing methods, and where it is fit for purpose.

New assessments are only one part of a process in which teachers will lead the way forward, curriculum and learning will change when teachers' practices change, however, assessment, in its various forms, can have a significant impact on this change.

6 References

Bennett, R. (2002), *Inexorable and Inevitable: The Continuing Story of Technology and Assessment*, The Journal of Technology, Learning and Assessment, Vol. 1, No. 1, June 2002.

Berry, R. (2003), *Alternative Assessment and Assessment for Learning*, IAEA Annual Conference 2003, Manchester, UK.

Black, P., Harrison, C., Lee, C., Marshall, B., Wiliam, D. (2003), *Assessment for Learning, Putting it into Practice*, Open University Press.

Eckstein, M., and Noah, H. (1993), *Secondary School Examinations, International Perspectives on Policies and Practice*, Yale University Press.

Grima, G. and Ventura, F (2000), Malta: *Lessons from the Past, Directions for the Future*, ACEAB First International Conference, Mauritius.

Kane, M. T., Crooks, T. and Cohen, A. (1999), *Validating Measures of Performance*, Educational Measurement: Issues and Practice, 19 (2) 5-17.

Mansell, W. (2003), The Times Educational Supplement, News & opinion 13; Issue No: 4533

Puntis, A., Beedle, P. and Maughan, S. (2003), *Assessing 'New Educational Goals'*, IAEA Annual Conference 2003, Manchester, UK.

Raffan, J. (2000), *School-based Assessment: Principles and Practice*, ACEAB First International Conference, Mauritius.

Russell, M. and Plati, T. (2001) *Effects of computer versus paper administration of a state-mandated writing assessment.* TC Record. Org.
http://www.tcrecord.org/content.asp?contentID=10709

VD Merwe, I. F. J. (2000), Namibia: SBA, *The Namibian Experience*, ACEAB First International Conference, Mauritius.

Wood, R. (1991), *Assessment and Testing, A Survey of Research*, Cambridge University Press.