Introduction of a range of computer-based objective tests in the examination of Genetics in first year Biology.

Sutcliffe, R G, Leonard, E M, Tierney, A, Howe, C W, Reid, I, (Institute of Biomedical and Life Sciences, University of Glasgow) **Goodwin, S T,** (Department of Computing Science, University of Glasgow), **Mackenzie, D M** (Division of Earth Sciences, University of Derby).

Address for correspondence: Division of Molecular Genetics, Institute of Biomedical and Life Sciences, University of Glasgow, 56 Dumbarton Road, G116NU

Abstract

The IBLS Level 1 Biology class of 750 provides opportunity for large-scale investigations of methods of assessment. Originally the examination consisted of multiple choice questions and several items of written work. Increased enrolment and pressure on staff time has necessitated reduction of the written component, though this has not been abolished. In January 1999, the end-of-module assessment in Genetics consisted of a multiple choice paper and two relatively new components: a set TRIADS-based questions, and optically marked sequencing tests supported by a new analysis programme. The MCQ and sequencing tests were held at a single sitting in an examination hall. The TRIADS test was invigilated and was run successively for groups of 50 students over the course of four days.

TRIADS questions were in text entry and drag-and-drop mode, and were applied to basic definitions, solving problems in inheritance, and to the comprehension of the chi^2 test. Student responses to one text-entry question involved a degree of explanation that precluded computer-marking; this text was exported for staff to mark 'by hand'. For the sequencing tests, a series of statements were provided in random order, on a stated topic. Some distracters were included, and students were asked to select correct statements and place them in the most appropriate order. The test was delivered summatively using OMR forms and a program was written to calculate marks.

We describe the introduction of TRIADS to the class, and the preparation and operation of the assessment and marking procedures. The outcomes for the different styles of question were interesting and provided insight into the limitations of multiple choice questions, compared with question styles that call for the co-ordinated analysis of several pieces of information. Despite the additional demands they make on students, these more demanding questions evoked no complaint from the class.

The Institute of Biomedical and Life Sciences of the University of Glasgow offers a wide range of modularised courses at Levels 1 and 2. Four Biology modules are offered to Level 1 students (n=750). Modules A and B are the subject of this report, and they are run in parallel during the first half of the year. Module A, *Plant Science and Biotechnology*, is examined through essay questions, short notes and MCQs. Module B is *Molecules, Cells and Genes*, which is examined through essays, short notes and problem-solving in Genetics. In previous years, problem solving in genetics was hand-written by students, and marked by staff in the normal way. This session we created a Genetics examination in TRIADS, and a new OMR-based sequencing test. The performance of these tests was compared with the other examination modes in Modules A and B.

TRIADS

Students were introduced to the TRIADS system in Week 6 of the course, during a laboratory session, and TRIADS material with feedback was made available to students for later practice. The TRIADS examination went through a total of three versions and was formally trialled on two occasions with groups of post-graduate laboratory demonstrators. TRIADS questions were in text entry, drag-and-drop and point-and-click modes. There were nine questions in all: the first three related to terminology and to problems on the inheritance of sex-linked characteristics in *Drosophila*; these were followed by three sets of pedigree problems. The final three questions addressed the calculation and interpretation of the chi^2 statistic. Of these, the final question required several sentences of text-entry by the students to explain the meaning and use of p values in the chi^2 test. This text was exported as typescript for staff to mark 'by hand'. The full examination took 45 minutes.

The final, packaged, version of the test was mounted on the server and the configuration process for recording of answers carried out on the morning of the first run of the test. The server operates 47 PCs in two rooms. In addition, 4 non-networked machines were available; these were separately set up and configured. A paper version of the test was available as a back-up and in addition there was a conventional test paper available in case of system failure. Students sat the test in groups of 40-50, in 45 minute sessions, repeated morning and afternoon for four days. Between the sessions, the examination and answer files were removed from the server, backed up, collated, and stored on another drive. The complete set of results files was available approximately 2 hours after the departure of the last student. Marking of the final question took academic staff approximately 45 seconds per answer.

In all, 708 students completed the computer-based test. Analysis was carried out on the results for 692 students: all of these had completed all parts of the assessment for this module and also for Module A (which runs concurrently). Correlations were calculated for a range of components, and three-dimensional graphical plots constructed.

Performance of TRIADS questions

No formal evaluation is available at this stage, although Dr Erica McAteer of the University's Teaching and Learning Service attended one of the demonstrator trials as evaluator. Student reaction to the Week 6 practice test was favourable, as was that of the demonstrators who took part in the trial tests. At the time of going to press, no negative comments have been received from the class.

The first analytical objective was to determine whether the students performed the test to a standard that was consistent with their performance in other examination formats in Modules A and B. Figure 1 shows the three-dimensional frequency distribution of marks from the TRIADS test plotted against the marks from the MCQs for the genetics component of the examination ($r = 0.46$).

Figures 2 and 3 show plots of the TRIADS data against marks for the essay and marks for the MCQ test of Module A ($r = 0.27$ and $r = 0.51$, respectively).

The correlations between the main components of the examinations for Modules A and B are summarised in Table 1.

Table 1 Correlation coefficients between the main components of the examination system for Modules A and B.

	Module B MCQ test	Module A MCQ test	Module A Essay	Module B TRIADS test
Module B MCQ test	1.0	0.47	0.21	0.46
Module A MCQ test		1.0	0.32	0.51
Module A Essay			1.0	0.27
Module B TRIADS test				1.0

Figure 1
The relationship between the marks for the TRIADS examination and the Module B MCQ test.

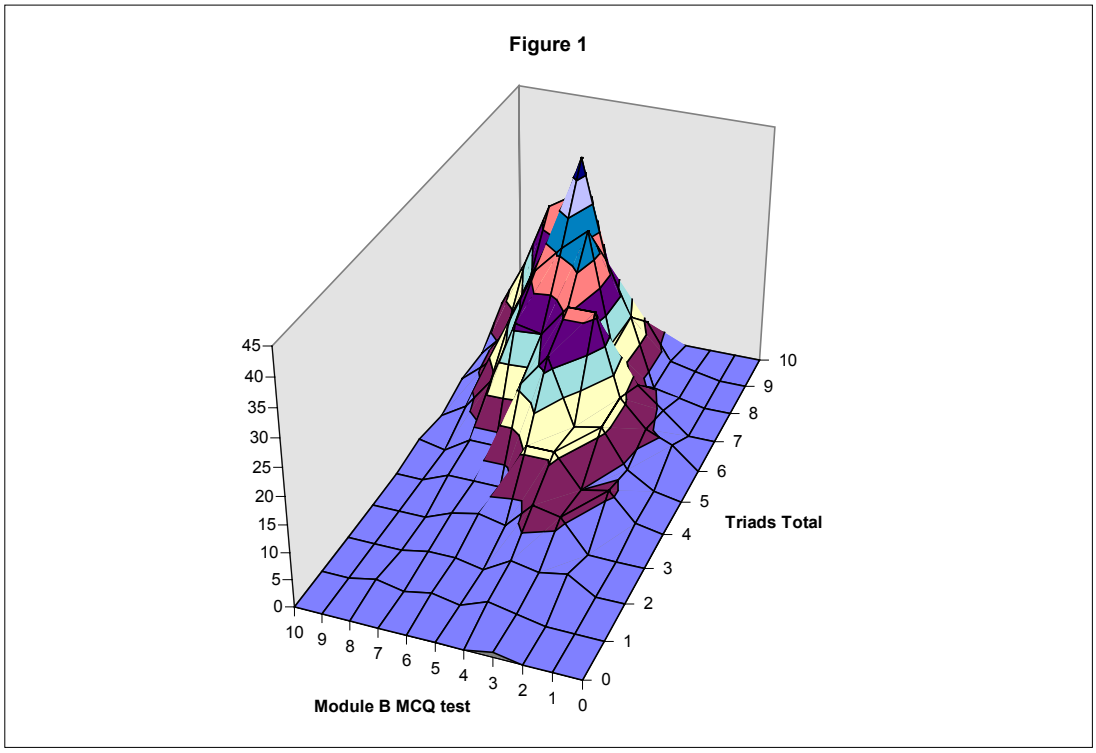


Figure 2
The relationship between the marks for the TRIADS examination and the Module A MCQ test.

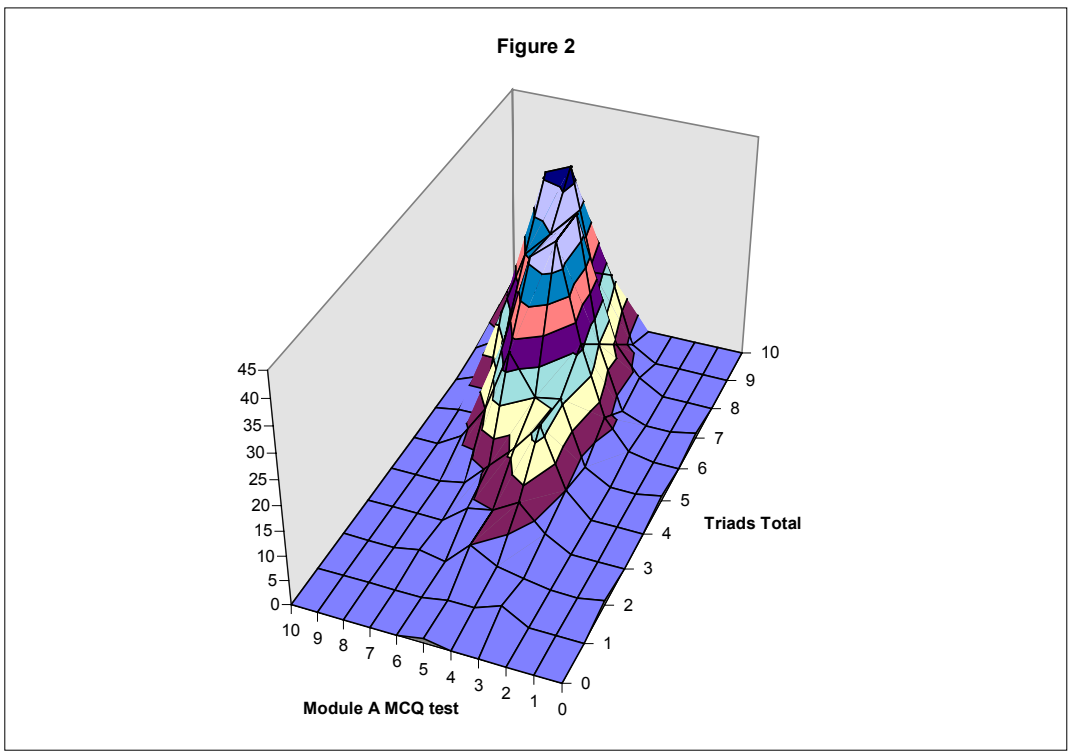
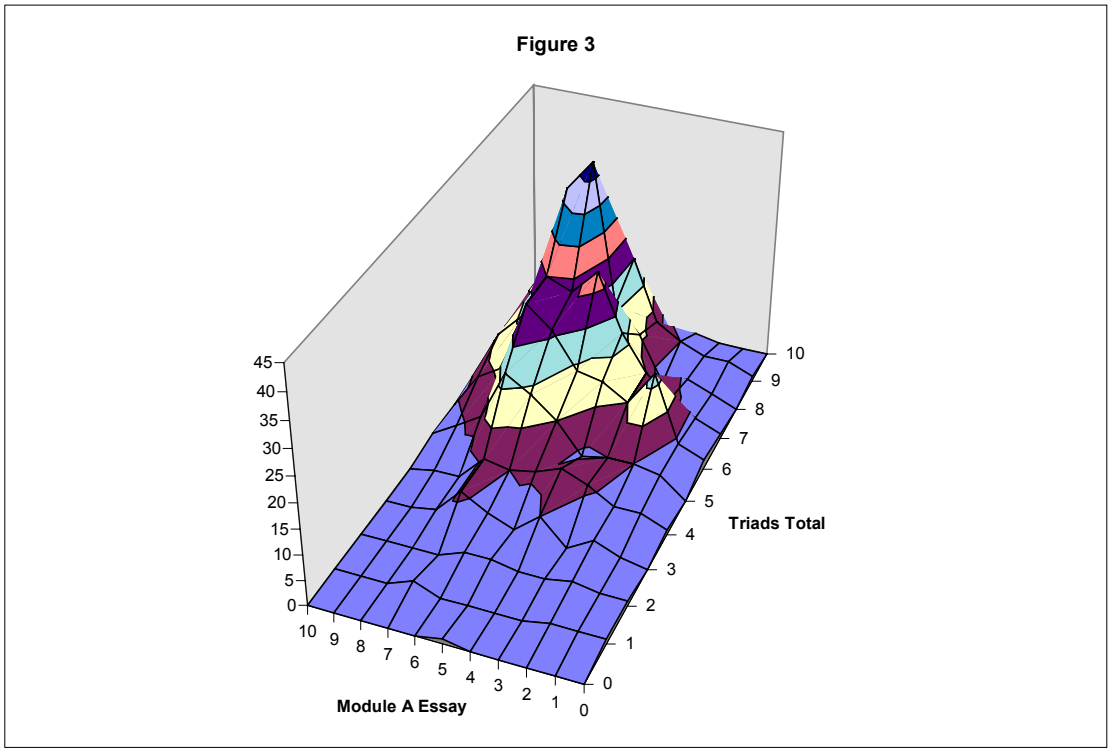


Figure 3
The relationship between the marks for the TRIADS examination and the Module A Essay



The sequencing test

The sequencing test was served to the class as a paper-based question in a formal examination setting. The question is shown in Table 2. It consisted of eight items relevant to the subject of haemophilia and four items that were irrelevant distracters. These items were presented on the question paper in a randomised order. Students were asked to select the relevant statements and arrange them in order from genotype to phenotype, i.e. in the order of events that lead to the development of the disease. Students entered their answers on a specially-designed optical mark reader sheet, and the answers were analysed with a purpose-designed programme. The order of correctly included items was assessed on a mark scale of 0-90; items erroneously included in the answer attracted the penalties shown in the far right column of Table 2. The resulting marks ranged between -47 to +90, and were normalised on a linear scale of 0 - 5 (less than -5 = 0 marks; 42 = 2.5 marks; 90 = 5.0 marks). Of the candidates, 14% scored over 80%; 39% scored >49% and 22% failed to score.

Table 2 The Sequencing Question.

Sequencing question *This question is about haemophilia. Select the appropriate statements from the list below and place them in the best order, from genotype to phenotype*			
Position of item in Question *	Question Items	Correct order of Inclusion in answer	
12	X chromosome	1	
5	Mutation	1	
3	Altered primary structure of polypeptide	2	
8	Defective function of factor VIII	3	
11	Prolonged clotting time	4	
6	Reduced wound healing	5	
1	Infection	5	Penalty points
9	Premature death	6	
2	Red cells lyse in blood vessels	Exclude	-5
7	Red cells sickle	Exclude	-31
10	Autosome	Exclude	-10
4	Haemoglobin S	Exclude	-31

- For the convenience of the reader, items are presented in order of biological action, with distracters at the end. The actual order of the items in the question is indicated by the numbers 1-12.

Table 3 A cognate multiple choice question

Multiple choice question	
Is sickle cell anaemia due to a defect in:	Percent selected
A. haemoglobin beta chain (correct answer)	47
B. haemoglobin alpha chain	20
C. blood clotting factor VIII	21
D. the MN blood group	12

Performance of the sequencing test

We wanted to compare the performance of students on the sequencing test with their scores from a multiple choice question on a similar topic. Since the sequencing test was about haemophilia (with distracters from sickle cell anaemia), we set the multiple choice question on sickle cell anaemia (with distracters from haemophilia and a blood group). The question is shown in Table 3, together with the frequency of selection of the alternative answers. The question was correctly answered by 47% of students, and 33% of students selected answers C or D, which where irrelevant to the structure and function of haemoglobin.

Table 4 shows the percentage of students who selected each multiple choice response, arranged in five ranks according to the mark attained for the sequencing question. There is a clear relationship between the results for the two questions. Those who obtained 3.7 - 5 marks on the sequencing test had the highest rate of selection of β-globin for the MCQ (60%). This percentage declined as the mark attained on the sequencing test declined. Of the students who scored 3.7 - 5 in the sequencing test, 11% selected the MCQ distracter "Factor VIII". The percentage selecting "Factor VIII" rose as the score attained for the sequencing question declined. No such trend was found for the MCQ responses for α-globin and the MN blood group. We believe this indicates that there is a significant group of students who know that sickle cell anaemia is a defect of haemoglobin function, but who cannot remember which polypeptide is defective. Such students have valuable partial knowledge. The small number of responses for the MN blood group indicate that the rate of 'pure guessing' was small. These observations give strong support for the concept that, though this sequencing test was quite severe, it is measuring academic ability in students.

Table 4 Performance on the multiple choice question, for students ranked by their performance in the sequencing test.

Sequencing question Mark Attained	Multiple Choice Question: Percentage of students selecting each response				Number of students
	β-globin	α-globin	factor VIII	MN blood group	
5.0 - 3.7	60	25	11	3	118
3.65 - 2.5	50	29	13	8	142
2.45 - 1.01	55	13	17	16	128
0.1 - 1.0	40	20	24	16	134
zero	44	19	31	6	154

Sequencing answers that scored very low marks were characterised by the presence of mutually exclusive items. For example, though 94% of students who scored 1 mark or less correctly included the item 'prolonged clotting time', 76% of them also included the distracters "haemoglobin S" and "red cells sickle". The selection of mutually exclusive items indicates that the students did not simply confuse the names for the two diseases. They seem to have failed to relate items to each other, or to have recognised that the items fall into two mutually exclusive groups.

We plan to investigate whether these response patterns are common in sequencing questions, or if they are peculiar to the present pair of questions. However, we know that level 1 students commonly check their factual knowledge against banks of MCQ questions, and we suspect that they may be failing to recognise that their learning should also focus on the inter-relationship of factual material. Students are not well practised in this type of puzzle-solving, and integrative learning is intrinsically more demanding and requires specific support. TRIADS lends itself to the further investigation of these issues.

Further comparisons between the tests

The sequencing test requires higher integrative skills, and this may explain the lack of an association between the results for the sequencing test and the results for the MCQ test for Module B (Figure 4). Figure 5 shows the relationship between the marks for the sequencing test and for the Essay (Module A). There is no great association; however, students who performed well in the sequencing test also performed well in the essay. It remains to be seen whether increased practice with sequencing tests (afforded by the availability of TRIADS) will improve student ability to plan and write essays.

Figure 4
The relationship between the marks for the sequencing test and the Module B MCQ test.

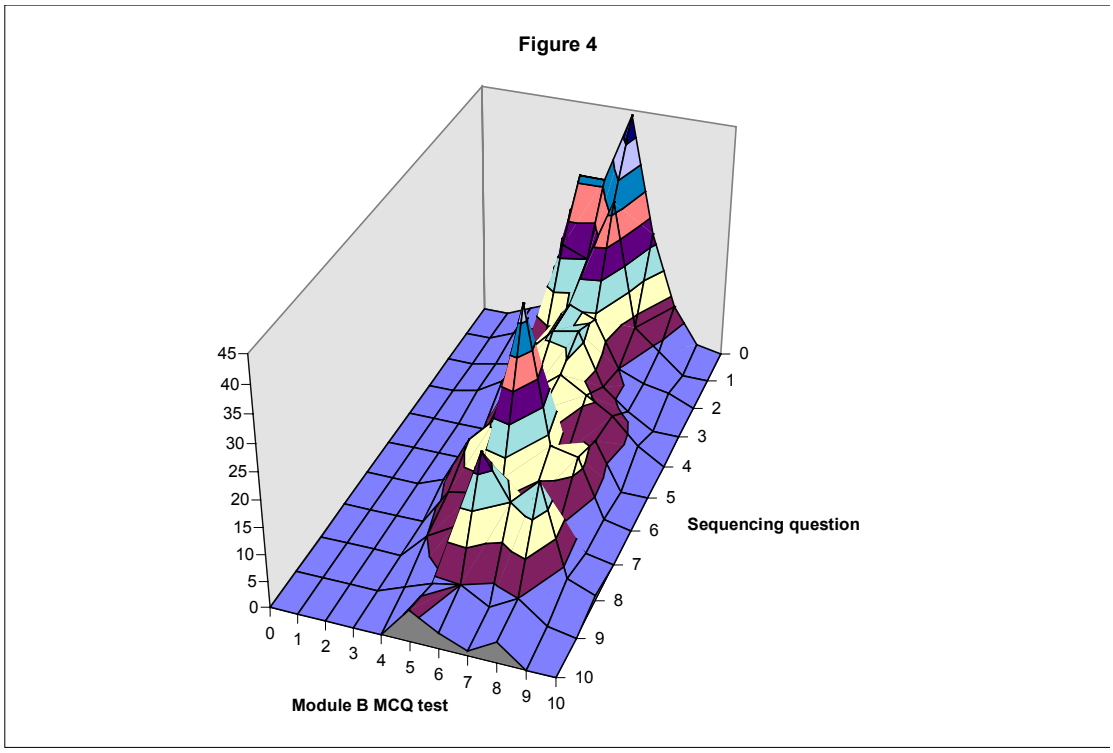
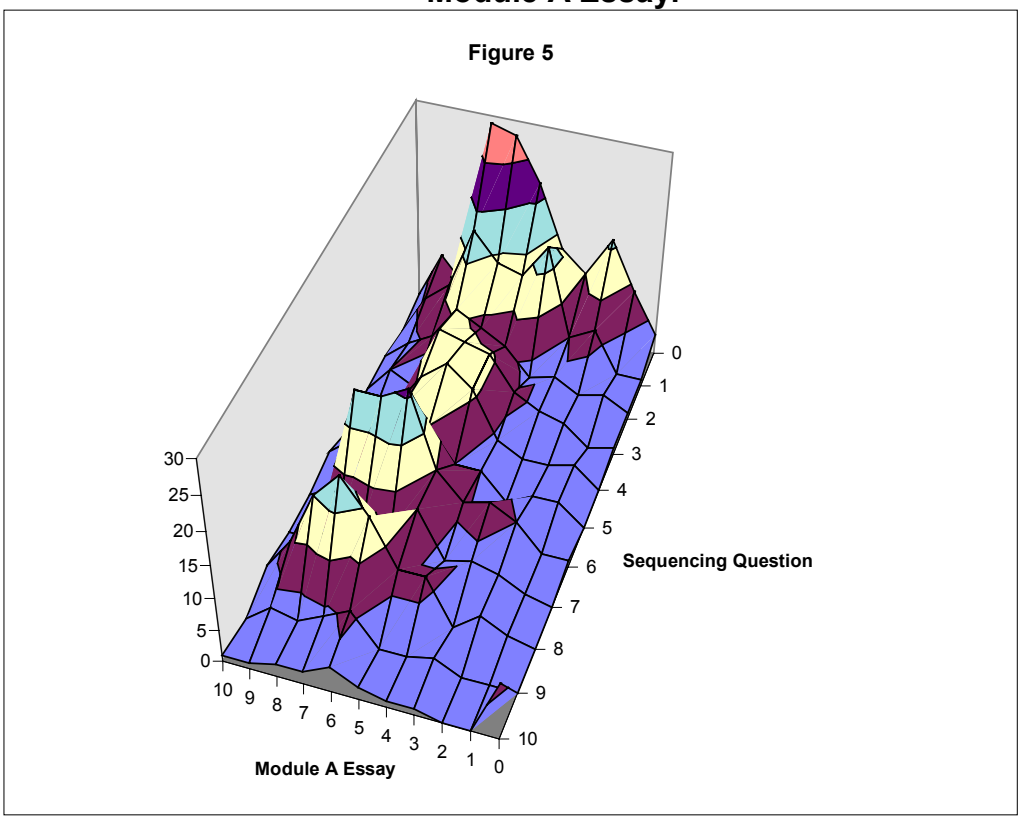


Figure 5
The relationship between the marks for the sequencing test and the Module A Essay.



Retrospect

For the first summative computer-based assessment with this course, and early in our experience with the TRIADS package, we feel that the computer-based part of this test was remarkably successful and trouble-free. Most problems were of a nature which had been foreseen and were dealt with by our protocol and back-up tests. Other minor difficulties should decrease as staff and students become familiar with the system.