*Acta Academica 2006 38(1): 233-252*

*Nebo Jovanovic & Abduraghiem Latief*

# Using multiple-choice questions, short questions and exercises to assess student performance

First submission: August 2004

Multiple-choice questions, short questions and exercises were used to assess the performance of students in environmental and water sciences at the University of the Western Cape (Bellville, South Africa). The objectives were to evaluate and compare these three assessment methods, and to indicate which were particularly appropriate for the various modules and academic years. Altogether 136 test and examination papers were marked. Students generally performed better in multiple-choice questions compared to short questions, while their performance on short exercises was the poorest. No significant difference in performance was found between the academic years. Multiple-choice questions are suitable for large classes, but short questions and exercises are recommended as they facilitate the assessment of conceptual knowledge and practical problem-solving skills.

## Meerkeusevrae, kortvrae en -oefeninge om studenteprestasie te evalueer

Meerkeusevrae, kortvrae en kortoefeninge is gebruik om die prestasie van studente in omgewings- en waterwetenskappe aan die Universiteit van die Wes-Kaap te evalueer. Die drie evalueringsmetodes is vergelyk om te bepaal watter evalueringsmetode geskik is vir die verskillende modules en akademiese jaarvlakke. 'n Totaal van 136 toets- en eksamenvraestelle is nagesien. Studente het beter gevaar met die meerkeusevrae as met die kortvrae, terwyl die kortoefeninge die swakste was. Daar was geen betekenisvolle verskille in die uitslag oor die akademiese jaargroepe nie. Meerkeusevrae is meer geskik vir groter klasse, maar kortvrae en kortoefeninge vergemaklik die evaluering van konseptuele kennis en praktiese vaardighede.

*Dr N Z Jovanovic, Dept of Earth Sciences, & Mr A Latief, Dept of Statistics, University of the Western Cape, Private Bag X17, Bellville 7535; E-mail: njovanovic@uwc.ac.za & alatief@uwc.ac.za*

In the last 20 years, there has been an increased interest in the different types of assessment methods (such as multiple-choice questions, short questions and exercises) used in higher education. Assessment is defined as a multi-dimensional process of judging individual students in action (Heywood 2000: 13). This process makes use of different forms of assessment, such as tests and examinations. It is widely recognised that assessment has a powerful influence on university learning because the primary goal of most university students is to pass examinations. A framework for the selection of suitable assessment methods was suggested by Ruddock (1981: 36-44). The choice of assessment methods depends on the desired outcomes and the level of educational system, institution, curriculum or lesson (module, course or syllabus). The teaching process then needs to be set up according to the objectives of the course, while the teaching results provide feedback to the primary objectives of education. In this way, feedback is provided to both teachers and students on improvements that could be introduced in a course (Marzano *et al* 1993: 1-5).

Heywood (2000: 316-72) reviewed a large number of assessment methods. In particular, short questions and multiple-choice questions are commonly used. Short questions require answers not longer than one or two paragraphs, while multiple-choice questions require the student to select one or more answers from a number of alternatives that are either right or wrong. Freeman *et al* (Heywood 2000: 349) assessed trainee graduate medical practitioners by making use of multiple-essay questions, an extended form of multiple-choice questions, where the possible answers offered various prognoses in medical cases. They found that trainees performed better on multiple-essay than in multiple-choice questions. Beullens *et al* (2002: 390-5) used between seven and 26 possible choices in multiple-choice questions to test medical students' knowledge. Shaha (1984: 469-81) reported that high school students performed better on matching items than on multiple-choice questions.

One of the major developments in higher education is the move towards the assessment of skills, competencies and abilities rather than knowledge only (Gibbs 1995: 63-72). This is driven by the market demand for professionals capable of tackling practical problems in the industrial, consulting, government and academic environments. Numerical exercises, where students are required to demonstrate problem-

solving skills are very suitable for science, engineering and technology subjects. Therefore, the use of a specific assessment method depends mainly on what outcome of learning needs to be assessed (Miller 1999: 149-54). For example, essay-type questions are more useful for assessing conceptual understanding, while multiple-choice questions are appropriate for testing factual knowledge (Miller *et al* 2000: 166-73). Other factors may also play a role in the choice of assessment method. For example, multiple-choice questions are easier and more consistent to mark (Bak 1990: 103-8; Borst 1990: 71-8) and computerise (Carbone *et al* 2000: 119-25; Lee & Weerakoon 2001: 152-7), while essays, short questions and exercises involve a greater degree of subjectivity on the part of lecturers (both in mark allocation and marking) (Miller 1996: 13-24; Reed 2003: 15-28). However, assessment by means of multiple-choice questions may permit students to guess the answer. The need to eliminate gueswork and the appropriateness of using four-option rather than three-option multiple-choice questions were discussed by Landrum & Cashin (1993: 771-8).

In the Department of Earth Sciences at the University of the Western Cape, a framework has been established for the three-year undergraduate course in environmental and water sciences. Based on this framework, first-year students are taught the basics of environmental and water sciences. Second-year students are trained in theory and modelling, with the aim of applying theoretical knowledge and mathematical skills to situations similar to those they will experience in life. Third-year coursework includes extensive practical experience by means of field visits, case studies and applied problems. The students' performance is assessed using multiple-objective examination (multiple-strategy assessment), defined as a comprehensive examination including subtests focused on well-defined domains of knowledge and skill (Heywood 1978: 216-34). The aim of multiple-objective examination is also to ensure equity and fairness of assessment (Paxton 2000: 109-19). For example, essays favour deep approaches to learning, focusing on meaning and understanding, while multiple-choice questions may favour more superficial approaches based on recall and reproduction (Scouller 1998: 453-72).

The general aim of this study was to compare three assessment methods (multiple-choice questions, short questions and exercises) in order

to recommend a standardised system for the assessment of students in the environmental and water sciences programme. The objective was to use the students' performance in terms of basic knowledge, conceptual understanding and problem-solving ability to determine which assessment methods are the most appropriate to the various modules and academic years. Specific objectives included the comparison of student performance across multiple-choice questions, short questions and exercises, and across academic years.

## 1. Materials and methods

The results of four one-term (half-semester) modules were investigated. The modules were EWS131 (Introduction to Environmental Science), taught to first-year students; EWS222 (Applied Soil Science), delivered to second-year students; ESC331 (Critical Resources of South Africa) and ESC336 (Applied Environmental Problems), presented to third-year (final-year) students. The EWS131 module comprised three one-hour lectures, one tutorial of one hour and one three-hour practical session per week. The EWS222 module was presented in three one-hour lectures and two three-hour practical sessions in the laboratory per week. The ESC331 and ESC336 modules involved three one-hour lectures and two three-hour practical sessions per week. The students were given the module descriptor (the objectives and outcomes of the course), the list of study material, the schedule of activities, the forms of assessment (test, assignment, examination and supplementary examination), the assessment criteria, the assessment standards and the marking system.

All lectures were presented with the aid of multi-media slide shows that included text, photos, video clips and sketches. The use of any particular medium depended on its appropriateness in communicating concepts in an easy and understandable way. Heap *et al* (1994: 649-57), as well as Sajaniemi & Kuittinen (1999: 269-84), proved that text colours and font style (especially bold face) can be used to direct learners' attention to different parts of the text and to the structure of the subject. Electronic formats are very suitable for highlighting relevant concepts. The technique of highlighting key concepts with different text colours and font styles was therefore widely used in the slide shows. The presentations were also made available to the students as lecture notes.

Forms of assessment included tests, assignments, examinations and supplementary examinations. Tests and assignments are part of the continuous assessment programme at the University of the Western Cape, and helped to enforce revision of course material by students during the term. The pass mark for each module was 50%. Students who did not pass the examination, but had a final mark of 45% wrote supplementary examinations. The final marks for all modules were calculated as a combination of coursework marks (40% of the final mark) and examination/supplementary examination marks (60% of the final mark).

The desired outcome of the practical sessions of the EWS131 module was to teach students how to report on technical topics and writing skills. This was done through two workshops on writing skills and assignments on technical topics. Continuous assessment was therefore applied through assignments, and assignment marks comprised the coursework mark. In the marking of the assignments particular attention was paid to structure, technical presentation and editing. After all the assignments had been marked, two practical sessions were devoted to discussion. No test was scheduled for EWS131.

The practical sessions for EWS222 included laboratory experiments that required reporting. The coursework mark was a combination of marks obtained from these reports on laboratory experiments and one test. The practical sessions for ESC331 and ESC336 included field trips to industrial, waste treatment and nature conservation sites. The students were required to hand in assignments on technical topics related to thse visits. For both ESC331 and ESC336, the coursework mark comprised a combination of the marks obtained from the assignments and one test. (No supplementary examinations were required for EWS222 and ESC331.)

Given the objectives of this study, only tests and examinations/supplementary examinations were taken into account. These will be discussed in detail below. Assignments will not be discussed further. The purpose of tests was mainly to help the students and the lecturers to recognise gaps in individuals' knowledge or understanding, and then to rectify possible weaknesses (Barrass 2002: 174-5). The tests were written at approximately mid-term and had an additional purpose of familiarising the students with the type of questions they could expect in the examination.

The test and examination papers were designed to assess students' performance in terms of basic knowledge, understanding of concepts and the ability to solve practical problems, rather than to measure their ability to memorise. The papers included a combination of multiple-choice questions, short questions and exercises. These were formulated in such a way as to cover the full module programme in a balanced way, and ordered in the sequence in which topics had been covered. Table 1 summarises the number of multiple-choice questions, short questions and exercises for each test and examination paper, as well as the marks allocated. Particular care was taken to ensure the even distribution of marks among multiple-choice questions, short questions and exercises in each paper, as one of the specific objectives of the study was the comparison of students' performance in terms of different assessment methods. However, this was not always feasible because multiple-choice questions lend themselves better to some topics than to others. The students were required to read all questions carefully. The maximum points for each question were indicated on the papers. The time allocated was deemed to be sufficient to answer all the questions (one hour for tests, and one-and-a-half hours for examinations). The total marks allocated to tests varied between 40 and 52, and to examination papers between 60 and 74 (Table 1).

The purpose of multiple-choice questions was to assess students' knowledge of basic facts and their ability to recall knowledge at relatively short notice. Students were required to give a quick answer, but also to have a thorough understanding of the question. Multiple-choice questions also helped to assess students' powers of deduction, for example, in eliminating wrong answers (Question 1, Table 2). Other examples of multiple-choice questions were conversions (Question 2, Table 2) as well as knowledge of the terminology and principles used in technologies (Question 3, Table 2). The assumption was that all multiple-choice questions were equally difficult and they were therefore allocated equal marks.

Multiple-choice questions are very suitable for testing knowledge recall, but they do not stimulate reasoning. Short questions and exercises were therefore included in the test and examination papers to create a multiple-strategy assessment. The purpose of short questions was to assess the students' ability to understand and report on concepts, scientific laws, hypotheses, theories and principles. For example, short

questions were related to assessment of knowledge of concepts (Question 4, Table 2), description of processes with visual aid (Question 5, Table 2), knowledge of scientific equipment (Question 6, Table 2) as well as the advantages and disadvantages of human activities for the environment (Question 7, Table 2). The lecturer allocated the marks for short questions subjectively, according to the level of difficulty.

The purpose of short exercises was to assess students' ability to apply knowledge and skills to real-world problems. The assessment of problem-solving skills was consistent with the topics for which the students were trained. For example, numerical exercises were related to practical applications (Question 8, Table 2) or theoretical knowledge (Question 9, Table 2). Students' knowledge of important chemical reactions taught in the course was also tested by this method (Question 10, Table 2). The lecturer allocated the marks for numerical exercises and chemical reactions subjectively, according to their level of difficulty.

In general, marks were awarded for the relevance, completeness and correctness of answers (Barrass 2002: 177-86). Precise, concise answers were required to the short questions and exercises, where students were required to keep to the point in order to demonstrate their critical thinking and their ability to discriminate for relevance. Marks were awarded for partially answered short questions and exercises. Students were required to show numerical calculations in the exercises in order to confirm their understanding and facilitate the marking of partial answers. No bonus marks were awarded for answering more than was required. Although the development of writing skills had been a major focal point of the course, no marks were deducted for poor writing or editing errors, as these were specifically dealt with in the practical sessions of EWS131.

The measurements of the three assessment methods were represented in percentages for comparison purposes. The statistical analyses were done to compare the performances of students on multiple-choice questions, short questions and exercises, both within modules and across academic years. The SAS statistical package was used to perform all the analyses. Due to the relatively small size of some of the samples, normality distribution tests were first run. These tests indicated that some sample data were not normally distributed. The statistical test used to compare students' performances on the three different assess-

ment methods within modules was the Kruskal-Wallis test for two or more samples (Van den Honert 1999: 256). In addition, the Kruskal-Wallis test was used to indicate which pairs might have caused differences to be significant. A 5% significance level applied to all tests.

Students' performances were also compared across academic years. Here again the Kruskal-Wallis test was used. Standard *z*-scores were used to compare average assessment marks expressed as percentages across the three academic years (Croucher 2004: 206). For the comparison of students' performances on multiple-choice questions, short questions and exercises across the academic years, the Kruskal-Wallis test was used for more than two samples as well as to compare pairs. Only examination marks were used, as these included independent groups of students. A 5% significance level applied to all tests. The examination format was relatively consistent across the academic years, although the study level, course content and examination papers were different. For the purpose of the Kruskal-Wallis test, it was assumed that the degree of difficulty of the modules and examination papers was equal across the academic years.

## 2. Results and discussion

Altogether 136 test and examination papers were marked, including 1872 multiple-choice questions, 506 short questions and 423 short exercises. The results and statistical analyses are shown in tabular format (Tables 3 to 6). All except Table 6 include test, examination and supplementary examination information for the sake of completeness. The students' results for each module and form of assessment (test, examination and supplementary examination) are set out in Table 3. The considerable difference between the maximum and the minimum marks reveals the wide range of students' backgrounds, in particular in EWS131 (Table 3). Some students needed the basics more than others and this had to be dealt with during individual tutorial classes. The average mark on the end-of-term examinations (60%) was higher than that on the mid-term tests (58%), while the average mark on the supplementary examinations was the lowest (52%) (see the last four rows of Table 3). This indicates that supplementary examinations are useful in increasing the pass rate, but not the overall average mark. The students obtained the highest marks in ESC331.

Overall, the average marks were 56% for first-years, 58% for second-years and 62% for third-years (Table 3). These averages were compared to the overall average (59%) and normalised for standard deviations to obtain standard $z$-scores. The standard $z$-scores were (56-59)/16 = -0.1875 for the first-years, (58-59)/16 = -0.0625 for the second-years and (62-59)/16 = 0.1875 for the third-years. These standard scores indicated that students obtained below-average scores in the first two years, and above-average scores in the third year, possibly due both to selection and to higher drop-out rates in the first and second years.

The results of the statistical analysis and the comparison between assessment methods (multiple-choice questions, short questions and exercises) are shown in Tables 4 and 5. Considerable variability in students' background was again evident from the large standard deviations (Table 4). In general, students performed better on multiple-choice questions (67%) than on short questions (59%), while their performance on short exercises was the poorest (41%). EWS222, where much emphasis had been placed on numerical exercises during lecturing, and the supplementary examination of ESC336, which included a very small sample, were the only exceptions. There was a significant difference between students' performance on each assessment method (Kruskal-Wallis test, $\chi^2$=68.6177, $p$-value < 0.0001) (Table 5). The differences between students' performances on each assessment method were confirmed by the total average score (last row in Table 4) and the overall statistical tests (last row in Table 5). This indicates that more attention should be paid to the development of students' problem-solving skills and to the practical application of theoretical problems.

The performances of students across all academic years was not significantly different (Kruskal-Wallis test, $\chi^2$=5.3917, $p$-value = 0.0675) (Table 6). A statistically significant improvement in students' performances was observed between first and third year for the short questions (Kruskal-Wallis test, $\chi^2$=11.5444, $p$-value = 0.0007). This indicated that the students' understanding of concepts and their reporting abilities had improved over the three years of study.

The good performance of students on multiple-choice questions could have been due to assessment errors such as gueswork. For example, all multiple-choice questions implied the choice of one out of three possible answers. This means that, where the answer was not known, there

was a 33.3% probability that a student could have answered correctly by guessing and a 66.7% probability of guessing wrongly. By assuming that 2/3 of not-known answers were guessed wrong and 1/3 guessed right, the total average mark of 67% (Table 4) can be recalculated: 67% - (100% - 67%)/2 = 51%. This figure lies between the total average marks obtained by students for short questions (59%) and for short exercises (41%). Higher and more stable marks for multiple-choice questions then for other assessment methods (open-book essays, for instance) are also recorded and discussed by Miller *et al* (2000: 166-73).

The results of tests and examinations were discussed in class in order to refine students' knowledge. In this way, the assessment of students' performances became an integral part of the coursework. Several reasons were identified for the under-performance of individual students. For example, absence from lectures was disadvantageous, as the main aim of direct contact was to outline and communicate important basics. In addition, more information was given in class than was included in the textbooks. Some students were complacent and achieved lower marks on the examination than on the mid-term test, while others did not read the assessment questions properly. Students' under-performance due to lack of focus is discussed by Selvaratnam & Mzibuko (1998: 42-6). Some students had difficulty in expressing themselves and answered too briefly, while others needed to improve their writing skills. Stress and anxiety could also have caused under-performance (Wilson & Louw 1997: 23-31), in particular for first-year students undergoing the transition from school to university. Individual support for such students is therefore essential. The forms of anxiety and the support tools that can be used are discussed in detail by Heywood (2000: 160-6).

## 3. Conclusions

The three methods of assessment employed in this study (multiple-choice questions, short questions and exercises) facilitated the identification of gaps in both teaching and learning. Multiple-strategy assessment is particularly suitable to groups of students from a wide range of backgrounds. The three assessment methods can also facilitate the evaluation of coursework and assist faculty management in making strategic decisions.

The continuous assessment programme, including mid-term tests, was beneficial in identifying students' needs and resolving problems timeously, as well as in suggesting possible improvements in coursework. Supplementary examinations were useful for increasing the pass rate, but not the overall average mark. The comparison between students' performances on multiple-choice questions, short questions and exercises indicated that more time should be spent on numerical problems and practical work. This could be extended to include exercises and problems with various possible solutions (Jovanovic & Annandale 2000: 15-22; Jovanovic *et al* 2000: 23-30) as well as team-work sessions. The statistical analysis of students' performances across academic years revealed how students' performance improved from the first year to the third year of study. However, students should be assessed consistently throughout their studies in order to measure the development of their skills over time.

The main limitation of the methodology utilised in this study was variability: in the population between academic years (as groups of people can never be homogeneous), between test and examination papers, and between test and examination questions. The effects of these variabilities cannot be easily overcome, but could be reduced by increasing the number of observations and by standardising tests and examinations.

A standardised system was recommended for the assessment of students in environmental and water sciences at the University of the Western Cape. Multiple-choice questions are suitable for large classes (especially for first-year students) as they facilitate and save time on marking. Short questions and exercises are recommended for small classes as they facilitate the assessment of conceptual knowledge and practical problem-solving skills.

Table 1: Number of multiple-choice questions, short questions and exercises for each test and examination paper, as well as marks allocated

| Module | Forms of assessment | Multiple-choice questions | | Short questions | | Short exercises | | Total marks allocated |
|---|---|---|---|---|---|---|---|---|
| | | Number | Marks allocated | Number | Marks allocated | Number | Marks allocated | |
| EWS131 | Examination | 16 | 32 | 4 | 16 | 3 | 12 | 60 |
| | Supplementary examination | 16 | 32 | 4 | 16 | 3 | 12 | 60 |
| EWS222 | Test | 12 | 24 | 2 | 14 | 2 | 14 | 52 |
| | Examination | 12 | 24 | 4 | 26 | 5 | 24 | 74 |
| ESC331 | Test | 12 | 24 | 2 | 8 | 2 | 8 | 40 |
| | Examination | 16 | 32 | 4 | 16 | 3 | 12 | 60 |
| ESC336 | Test | 8 | 16 | 3 | 12 | 3 | 12 | 40 |
| | Examination | 12 | 24 | 5 | 20 | 4 | 16 | 60 |
| | Supplementary examination | 12 | 24 | 5 | 20 | 4 | 16 | 60 |

Table 2: Examples of multiple-choice questions, short questions and short exercises used to assess students' performances

| Method | Question | Mark allocated | Module |
|---|---|---|---|
| Multiple -choice | 1. What is point source pollution? | 2 | EWS131 |
| | a) Pollution originating from dispersed sources | | |
| | b) Pollution originating from one source in a region | | |
| | c) Pollution originating from single, identifiable sources | | |
| | 2. 10 m of water column corresponds to: | 2 | EWS222 |
| | a) -1 J $kg^{-1}$ | | |
| | b) -100 J $kg^{-1}$ | | |
| | c) -1000 J $kg^{-1}$ | | |
| | 3. What is bioremediation? | 2 | ESC336 |
| | a) Technology for water purification or waste degradation making use of bacteria | | |
| | b) Organisms' response to gene mutation | | |
| | c) Organisms' response to changes in the environment | | |
| Short questions | 4. Explain in your own words the limiting factor principle. | 4 | EWS131 |
| | 5. Describe the components of the carbon cycle with the aid of a sketch. | 4 | EWS131 |
| | 6. Discuss briefly the principle of time domain reflectometry. | 6 | EWS222 |
| | 7. Discuss briefly at least 4 impacts of coal mining on the environment. | 4 | ESC331 |
| Short exercises | 8. If 50 kg of radioactive waste having a half-life of 200 years are stored, what will the mass of radioactive material be in 1,000 years? | 4 | EWS131 |
| | 9. What is the bulk density of two soils having porosity of 50% and 60%? | 6 | EWS222 |
| | 10. Write the chemical reaction of ozone and chlorine (ozone depletion). | 4 | ESC336 |

Table 3: Results: students' performances

| Module | Assessment form | Number of papers | Mark (%) | | | | | Pass rate (%) |
|---|---|---|---|---|---|---|---|---|
| | | | Average | Maximum | Mininum | Median | Standard deviation | |
| EWS131 | Examination | 54 | 57 | 87 | 17 | 58 | 16 | 67 |
| | Supplementary examination | 6 | 52 | 70 | 40 | 51 | 12 | 50 |
| | Total | 60 | 56 | 87 | 17 | 57 | 16 | 65 |
| | | | | | | | | |
| EWS222 | Test | 8 | 56 | 71 | 31 | 61 | 14 | 75 |
| | Examination | 9 | 59 | 83 | 32 | 61 | 17 | 78 |
| | Total | 17 | 58 | 83 | 31 | 61 | 16 | 76 |
| | | | | | | | | |
| ESC331 | Test | 13 | 68 | 85 | 48 | 70 | 13 | 92 |
| | Examination | 14 | 76 | 93 | 50 | 75 | 13 | 100 |
| | Total | 27 | 72 | 93 | 48 | 75 | 13 | 96 |
| | | | | | | | | |
| ESC336 | Test | 14 | 49 | 83 | 37 | 43 | 14 | 29 |
| | Examination | 15 | 58 | 88 | 35 | 58 | 14 | 80 |
| | Supplementary examination | 3 | 52 | 55 | 50 | 50 | 3 | 100 |
| | Total | 32 | 53 | 88 | 35 | 50 | 14 | 59 |
| | | | | | | | | |
| ESC300 (including ESC331 and ESC336) | | 59 | 62 | 93 | 35 | 62 | 16 | 76 |
| | | | | | | | | |
| Tests | | 35 | 58 | 85 | 31 | 58 | 15 | 63 |
| Examinations | | 92 | 60 | 93 | 17 | 62 | 17 | 75 |
| Supplementary examinations | | 9 | 52 | 70 | 40 | 50 | 10 | 67 |
| Total | | 136 | 59 | 93 | 17 | 59 | 16 | 71 |

Table 4: Marks awarded, average mark of students and standard deviation for each assessment method

| Module | Forms of assessment | Multiple-choice questions | | Short questions | | Short exercises | |
|---|---|---|---|---|---|---|---|
| | | Average (%) | Standard deviation (%) | Average (%) | Standard deviation (%) | Average (%) | Standard deviation (%) |
| EWS131 | Examination | 68 | 17 | 50 | 26 | 37 | 22 |
| | Supplementary examination | 66 | 15 | 43 | 18 | 26 | 8 |
| | | | | | | | |
| EWS222 | Test | 60 | 19 | 50 | 16 | 55 | 36 |
| | Examination | 67 | 19 | 57 | 21 | 54 | 23 |
| | | | | | | | |
| ESC331 | Test | 74 | 20 | 71 | 21 | 43 | 37 |
| | Examination | 81 | 12 | 85 | 13 | 51 | 35 |
| | | | | | | | |
| ESC336 | Test | 46 | 12 | 70 | 15 | 33 | 28 |
| | Examination | 67 | 17 | 58 | 21 | 43 | 18 |
| | Supplementary examination | 56 | 5 | 50 | 5 | 48 | 10 |
| | | | | | | | |
| | Total | 67 | 18 | 59 | 24 | 41 | 26 |

Table 5: Kruskal-Wallis statistical analysis of students' performances across assessment methods (multiple-choice questions MCQ, short questions SQ and short exercises EX) for each assessment

| Module | Forms of assessment | Number of observations | Statistical values | | | |
|---|---|---|---|---|---|---|
| | | | MCQ = SQ = EX | MCQ = SQ | MCQ = EX | SQ = EX |
| EWS131 | Examination | 54 | 40.4875 (<0.0001) | 12.6171 (0.0004) | 40.7758 (<0.0001) | 7.2732 (0.0070) |
| | Supplementary examination | 6 | 9.8768 (0.0025) | 4.0777 (0.0435) | 8.4857 (0.0036) | 2.1139 *(0.1460) |
| EWS222 | Test | 8 | 0.8476 *(0.6546) | N/A | N/A | N/A |
| | Examination | 9 | 1.8834 *(0.39) | N/A | N/A | N/A |
| ESC331 | Test | 13 | 6.3126 (0.0426) | 0.0667 *(0.7962) | 5.0418 (0.0247) | 4.3398 (0.0372) |
| | Examination | 14 | 7.6256 (0.0221) | 0.8692 *(0.3512) | 4.7988 (0.0285) | 5.8430 (0.0156) |
| ESC336 | Test | 14 | 17.6432 (<0.0001) | 12.9830 (0.0003) | 3.0645 *(0.0800) | 11.5538 (0.0007) |
| | Examination | 15 | 10.3502 (0.0057) | 1.7186 *(0.1899) | 10.0017 (0.0016) | 3.6608 *(0.0557) |
| | Supplementary examination | 3 | 2.5043 *(0.2859) | N/A | N/A | N/A |
| Overall | | 136 | 68.6177 (<0.0001) | 6.3303 (0.0119) | 65.5655 (<0.0001) | 30.9433 (<0.0001) |

Note: Values in brackets are *p*-values * denotes that the result is non-significant

Table 6: Kruskal-Wallis statistical analysis of students' performances in examinations on each assessment method (multiple-choice questions MCQ, short questions SQ and short exercises EX) across academic years

| Assessment method | Statistical values | | | |
|---|---|---|---|---|
| | EWS131 = EWS222 = ESC300[1] | EWS131 = EWS222 | EWS131 = ESC3001 | EWS222 = ESC3001 |
| MCQ | 2.0983 *(0.3502) | N/A | N/A | N/A |
| SQ | 11.5863 (0.0030) | 0.2425 *(0.6224) | 11.5444 (0.0007) | 2.3986 *(0.1214) |
| EX | 5.6812 *(0.0584) | N/A | N/A | N/A |
| Overall | 5.3917 *(0.0675) | N/A | N/A | N/A |

[1] ESC300 includes ESC331 and ESC336

Note: Values in brackets are *p*-values * denotes that the result is non-significant

# Bibliography


BAK N

1990. How to test insight and understanding of philosophical issues by means of multiple-choice questions. *South African Journal of Education* 10(2): 103-8.

BARRASS R

2002. *Study! A guide to effective learning, revision and examination techniques*. 2nd ed. London: Routledge.

BEULLENS J, B VAN DAMME, H JASPAERT & P J JANSSEN

2002. Are extended-matching multiple-choice items appropriate for a final test in medical education? *Medical Teacher* 24(4): 390-5.

BORST C B

1990. Meerkeusevrae: 2. *Progressio* 12(1): 71-8.

CARBONE A, P SCHENDZIELORZ & J D ZAKIS

2000. Electronic assessment and self-paced learning on the Web using a multiple-choice quiz generator. *International Journal of Electrical Engineering Education* 37(2): 119-25.

CROUCHER J S

2004. *Introductory Mathematics and Statistics for Business*. 4th ed. City: McGraw-Hill.

FREEMAN J, J ROBERTS, D METCALFE & V HILLER

1982. *The influence of trainers on trainees in general practice*. Occasional Paper, 21. London: Royal College of General Practitioners.

GIBBS G

1995. *Assessing student centred courses*. Oxford: Oxford Brookes University.

HEAP I M, L C BURRILL, S A DEWEY & G E MACDONALD

1994. Eradication of noxious visual aids: slides and overheads that work. *Weed Technology* 8: 649-57.

HEYWOOD J

1978. *Assessment in higher education*. Chichester: Wiley.

2000. *Assessment in higher education. Student learning, teaching, programmes and institutions*. London: Jessica Kingsley.

JOVANOVIC N Z & J G ANNANDALE

2000. SWB (Soil Water Balance): a computer tool for teaching future irrigation managers. *Journal of Natural Resources and Life Science Education* 29: 15-22.

JOVANOVIC N Z, J G ANNANDALE & P S HAMMES

2000. Teaching crop physiology with the Soil Water Balance (SWB) model. *Journal of Natural Resources and Life Science Education* 29: 23-30.

LANDRUM E R & J R CASHIN
1993. More evidence in favor of three-option multiple-choice tests. *Educational and Psychological Measurement* 53(3): 771-8.

LEE G & P WEERAKOON
2001. The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple-choice tests. *Medical Teacher* 23(2): 152-7.

MARZANO R J, D PICKERING & J MCTIGHE
1993. *Assessing student outcomes. Performance assessment using the dimensions of learning model*. Alexandria, VA: Association for Supervision and Curriculum Development.

MILLER R
1996. Mark my words. Part 1, Teachers. *South African Journal of Higher Education* 10(2): 13-24.

1999. The half-life of semester-long learning. *South African Journal of Higher Education* 13(3): 149-54.

MILLER R, J BRADBURY & G LEMMON
2000. Justifying means with ends: assessment and academic performance. *South African Journal of Higher Education* 14(1): 166-73.

PAXTON M
2000. A linguistic perspective on multiple-choice questioning. *Assessment and Evaluation in Higher Education* 25(2): 109-19.

REED Y
2003. (Un)reliable assessment: a case study. *Perspectives in Education* 21(1): 15-28.

RUDDOCK R
1981. *Evaluation: a consideration of principles and methods*. Manchester Monographs, 18. Manchester: University of Manchester. Dept. of Adult and Higher Education.

SAJANIEMI J & M KUITTINEN
1999. Three-level teaching material for computer-aided lecturing. *Computers and Education* 32(4): 269-84.

SCOULLER K M
1998. The influence of assessment methods on students' learning approaches: multiple-choice question examination versus assignment essay. *Higher Education* 35: 453-72.

SELVARATNAM M & B MZIBUKO
1998. Importance of focusing sharply on the goal for successful problem solving. *South African Journal of Chemistry* 51(1): 42-6.

SHAHA S H
1984. Matching tests: reduced anxiety and increased test-effectiveness. *Educational and Psychological Measurement* 47(4): 469-81.

VAN DEN HONERT R
1999. *Intermediate statistical methods*. 2nd ed. Cape Town: UCT Press.

WILSON D A B & G J LOUW
1997. Stress, anxiety and academic performance among medical students at the University of Cape Town. *Southern African Journal of Child and Adolescent Mental Health* 10(1): 23-31.