

# Teaching and Learning

beyond Formal Access

Assessment through  
the Looking Glass

HE Monitor No. 10  
August 2010

# Teaching and Learning beyond Formal Access

Assessment through  
the Looking Glass

HE Monitor No.10

August 2010

The South African Council on Higher Education (CHE) is an independent statutory body responsible for advising the Minister of Higher Education and Training on all higher education policy issues, and for quality assurance in higher education and training.

Published by the Council on Higher Education (CHE) in 2010

1 Quintin Brand Street  
Perseus Technopark  
Brummeria  
Pretoria  
South Africa  
+27 12 349 3840  
[www.che.ac.za](http://www.che.ac.za)

Produced on behalf of the CHE by Jacana Media

10 Orange Street  
Sunnyside  
Auckland Park 2092  
South Africa  
+27 11 628 3200  
[www.jacana.co.za](http://www.jacana.co.za)

© Council on Higher Education Pretoria, 2010

All rights reserved. Material from this publication may not be reproduced without the CHE's permission.

ISBN 978-1-919856-77-3

Set in Sabon 10.5/14pt  
Printed and bound by Paradigm Print, Johannesburg  
Job No. 001206

See a complete list of Jacana titles at [www.jacana.co.za](http://www.jacana.co.za)

# Contents

Preface.....	v
<i>Lis Lange</i>	
Chapter 1: Towards a more authentic approach to assessment: An innovative response to large class assessment.....	1
<i>Susan Benvenuti</i>	
Chapter 2: Using MCQ-based assessment to achieve validity, reliability and manageability in introductory level large class assessment.....	21
<i>Susan Benvenuti</i>	
Chapter 3: A major effect of a minor intervention: The use of a classification instrument to create awareness of assessment practices amongst lecturers in a first year biology course.....	35
<i>Elisabeth Brenner, Grace Moletsane and Marissa Rollnick</i>	
Chapter 4: Item analysis as a useful diagnostic tool in the design and analysis of multiple choice questions.....	51
<i>Marissa Rollnick, Elisabeth Brenner and Grace Moletsane</i>	
Chapter 5: Immediate response to mediated learning: The use of technology for continuous assessment in higher education.....	65
<i>Elisabeth Brenner and Yael Shalem</i>	
Chapter 6: Reading below the surface: Students' organisation of content and form.....	81
<i>Lynne Slonimsky and Yael Shalem</i>	
Contributors.....	110



# Preface

This issue of the *Higher Education Monitor* presents research done in the context of the Finland–South Africa Cooperation Agreement, which ran between 2005 and 2008 under the auspices of the Department of Education. The Department of Education approached the Council on Higher Education to participate in the cooperation agreement with a bid for funding in the area of quality assurance. The funds obtained were disbursed in two broad project areas. One focused on supporting the development of quality assurance systems in merged and historically disadvantaged institutions. The other focused on supporting projects geared to the improvement of teaching and learning at institutions which did not fall in either of the other two groups. Funds in the improvement of teaching and learning stream were allocated in the form of competitive grants.

Several universities applied for and obtained grants in this project area and all of them have produced interesting and innovative work. This publication is but one example of the kind of work produced at our universities in order to improve teaching and learning. We are hopeful that in the course of 2010–11, we will be able to publish other examples of relevant work in the area of teaching and learning done in contexts other than the Finland–South Africa Agreement.

This publication is organised in six chapters which reflect collaborative research done by specialists in education and academics responsible for teaching and learning of subjects offered at different academic units at the University of the Witwatersrand (Wits): the Faculty of Commerce, Law and Management, the Faculty of Science and the School of Education. Chapters One to Five focus on research conducted on actual assessment practices in large classes and the impact that the introduction of alternative approaches had in making assessment not only more valid, reliable, effective and manageable, but also in making assessment a part of students' learning. Chapter Six explores the manner in which assessment that is text-based (essays in this case) respond to the concept of assessment for learning. This chapter grapples with a range of issues about the acquisition of academic knowledge and the assumptions underpinning the essay questions and the answers provided by students.<sup>1</sup>

All the chapters in this publication present theory-informed research that could be transferred to other contexts, illuminate similar problems, or confirm variations on innovative practice. Above all, this publication confirms that teaching and learning, far from being common-sense activities, require specialist knowledge which is theoretically informed and technically supported. It confirms that the improvement of teaching and learning can take place in the interstices between specialist disciplinary knowledge, skills and values, and pedagogic discourse and practice, and that those interstices need to be purposefully looked for and opened up at institutional level.

From the point of view of higher education policy, this issue of the *Higher Education Monitor* is about student access and quality. Every piece of higher education policy since 1996 has set access as one of the most important goals for the higher education system in the democratic transition. Given

---

1 The original plan for the book included another chapter, 'Reading below the surface: Students' organisation of content and form' by Yael Shalem and Lynne Slonimsky. Fortunately for the authors, their chapter was accepted for publication in *The British Journal of Sociology of Education*, Volume 31, Number 6 (November 2010), which precludes its publication in this volume.

the history of South Africa, access means not only more students in higher education, but a student population that reflects the demography of the country. Important research on higher education has focused on the issue of access, indicating that whatever gains have been made in growing the size of the higher education system, our participation rates are unsatisfactory and our throughput and dropout rates a reason to wonder about the extent of the deracialisation of the higher education system and its overall effectiveness (Bunting, 1994; Cloete et al., 2002; CHE, 2004 and Scott et al., 2007). The Higher Education Quality Committee's (HEQC) institutional audits are showing that in the area of teaching and learning, growing access (undoubtedly a success in the first 15 years of democracy) has created a number of new challenges for higher education institutions, which range from the suitability of their educational infrastructure to the suitability of their pedagogic approaches. Yet, higher education institutions, their management and academic staff are not the only ones tested by the (insufficient) expansion of the higher education system. The 'problem of access' also questions the facile responses provided in the public (social and political) discourse, which often only sees access in terms of universities reluctant to fulfil their duties to society and underprepared black students who are not ready for university education.

The research gathered in this publication provides a sharp example of the two dimensions of access identified over 20 years ago by Wally Morrow and which constituted a constant focus of his professional reflection: formal access and epistemological access (Morrow, 2007). The inevitable consequence of expanding access and achieving greater equity in our society has been growth in the number of student enrolments. Large classes are the outcome of the commitment to formal access. This has meant that all higher education institutions, to a greater or lesser extent, have been confronted with lecturers' difficulties in teaching large classes, particularly at first-year level. In this context, frustration among students and lecturers is coupled with a wish for smaller classes in which teaching and learning is regarded as more effective. As Morrow noted, not only is class size unlikely to shrink, but also, all mass higher education systems have had to learn to teach large classes. Against this backdrop, he wondered why it was it so difficult for us in South Africa to think of large classes as a pedagogic problem to be solved and not as a developing country aberration and the first step in the dropping of standards. As the research presented here shows, it is possible to teach and assess large classes without dumbing down content and skills, but this requires concerted interventions. It requires academic learning and the support of subject specialists in order to think about the syllabus pedagogically. This, in turn, requires institutional awareness of the need to invest in educational infrastructure and human resources as well as of the need to acknowledge and value the demands that effective teaching makes of academics, including the professionalisation of teaching at higher education level. While the institutional audits conducted by the HEQC show instances of institutions struggling to rise to the occasion, they also show important examples of individual and institutional efforts to deal with large classes in innovative and professional manners. The research presented here is but one example of this.

The other area of concern in Morrow's writing is epistemological access, or 'access to the knowledge that universities distribute', and although this does overlap with the challenges posed by the teaching of large classes, it is not necessarily the same. Epistemological access is a political as well as an educational issue in that it turns the spotlight both on to unconscious and unquestioned processes of concept formation and knowledge acquisition, and on to the assumptions that inform the manner in which teaching at university level takes place. The construction of students as autonomous subjects who, by virtue of having access to higher education, are going to actualise their potential making the most of the opportunities offered to them, is common to many universities in the country. However, statistics on student success and interviews with students and lecturers show that the notion of the autonomous subject is not helping to provide a fruitful student experience at universities and gets in the way of helping students in the process of achieving epistemological access (Boughey, 2007; CHE, 2010). It is in relation to epistemological access where the greatest work remains to be done in the country; and this work, as most of the literature in the field suggests, involves a different

way of looking at teaching and learning at university level, greater resourcing, a refocus on the centrality of the curriculum and a greater understanding of the importance of research on teaching and learning.

As indicated earlier, this issue of the *Higher Education Monitor* is also about quality. In its Founding Document the HEQC stated its commitment to ‘a quality-driven higher education system that contributes to socio-economic development, social justice and innovative scholarship in South Africa’ (CHE, 2001:9). Only with the confluence of formal and epistemological access will higher education be able to realise its part in bringing about social justice to South Africa.

The promotion of an understanding of quality education that holds together equity and standards and the support of higher education institutions in their efforts to achieve greater equity and quality are an important part of the quality assurance work of the Council on Higher Education. It is hoped that this publication will be of use to academics, education specialists and higher education managers, that it will generate debate and, possibly, enthusiasm for reviewing and renewing teaching and learning practices.

Finally, I would like to thank Prof. Yael Shalem and her colleagues at Wits for their work on this publication. In them I thank all lecturers and managers at higher education institutions whose commitment to access continues unabated in the face of many systemic and institutional difficulties and whose work, we hope, will be the focus of further publications aimed at the promotion of quality teaching and learning.

*Dr Lis Lange*  
Executive Director  
Higher Education Quality Committee  
Pretoria  
March 2010

## References

- Boughey, C, 2007. Educational Development in South Africa: From Social Reproduction to Capitalist Expansion? *Higher Education Policy*, 20, 5–18.
- Bunting, I, 1994. *A Legacy of Inequality: Higher Education in South Africa*. Cape Town: University of Cape Town Press.
- Cloete, N, Fehnel, R, Maasen, P, Moja, T, Perold, H & Gibbon, T (Eds.), 2002. *Transformation in Higher Education: Global Pressures and Local Realities in South Africa*. Cape Town: Juta.
- Council on Higher Education, 2001. *Higher Education Committee Founding Document*. Pretoria: CHE.
- Council on Higher Education, 2004. *South African Education in the First Decade of Democracy*. Pretoria: CHE.
- Council on Higher Education, 2010. Access and Throughput in South African Higher Education: Three Case Studies. *Higher Education Monitor*, 9. Pretoria: CHE (forthcoming).
- Morrow, W, 2007. *Learning to Teach in South Africa*. Cape Town: HSRC Press.
- Scott, I, Yeld, N & Hendry, J, 2007. A Case for Improving Teaching and Learning in South African Higher Education. *Higher Education Monitor*, 6. Pretoria: CHE.



# Towards a more authentic approach to assessment

## An innovative response to large class assessment

*Susan Benvenuti*

### Introduction

Assessment continues to be a challenging balancing act, with concerns for validity and reliability on the one side of the scale and feasibility or manageability on the other. So despite the now widely held view that assessment tasks and practices determine (or at best strongly influence) what students learn and how they engage with course curricula (Biggs, 2003; Brown, Bull & Pendlebury, 1997; Brown & Knight, 1994; Boud, Cohen & Sampson, 1999; Entwistle & Entwistle, 1997; Rowntree, 1987; and others), when faced with classes of several hundred students and limited time and resources, educators invariably take refuge in the time-honoured practice of traditional timed assessments.

In this chapter we examine an innovative response taken by academics to the challenge of a semester-long introductory level course in information systems (IS) with enrollment numbers that have topped 900 students in the very recent past. The challenge was twofold: firstly, to design an assessment strategy that would allow for a more authentic engagement by the students with the broad range of knowledge, skills and values that were important outcomes of the course, and, secondly, to ensure that the implementation of this strategy provides a valid and reliable demonstration of student ability and achievement that is manageable and feasible for the academics, given the time and resource constraints.

### Background to the IS course

The first year IS course under discussion is an introductory course taken by virtually all BCom students at the university. At the start of this research it ran as a half course over a full year, but now runs over the period of a single semester, covering a wide range of theoretical topics relating to technology and its use in business settings, as well as requiring students to master some procedural knowledge by way of learning some widely used software products.

The course is a compulsory course for BCom graduates as it is intended to provide sufficient background in the IS discipline for all students intending to work in the business or commercial sphere. Regardless of whether students hope to major in marketing, finance, management, human resources, economics, etc. or go on to work in any other business-related field, they are expected to have sufficient exposure to enable them to both use, and make decisions relating to the use of,

information systems in their respective fields. By the end of this IS module, they should be expert users of information systems and demonstrate an understanding of how and why IS contributes to business success.

## The research project

When funding became available for the ‘Assessment for Learning’ research project, the IS course was put forward for consideration as it met the broad criteria of the project. With the eventual goal of rolling out the findings from the project into many of the first year courses across the university, it represented a compulsory, introductory course from the Commerce, Law and Management (CLM) Faculty, with large classes and staffing constraints, and ties with courses in the faculty with similar challenges. In fact, all first year courses in the faculty would have very similar issues in terms of student numbers and many of the other challenges identified below, precisely because the students taking those courses would be the same as those taking the IS course.

## Course challenges

The research project began by establishing the challenges facing the IS course and examining the current assessment practices. Some of the challenges identified were as follows:

- **Large classes:** At the time of the start of the research project, the IS course was heading into its first year as a compulsory course for the BCom degree, and was expected to have an enrolment of between 400 and 500 students in the following year. At the time of writing, three years later, the student enrolment for the year just completed was over 900 students.
- **Mixed educational backgrounds:** Linked to the larger class sizes is an increase in the diversity of students in a class, with a wide-ranging spectrum of ability, background and preparedness for tertiary education, and motivation (Gibbs & Jenkins, 1992, cited in Benvenuti & Cohen, 2008).
- **Academic literacy and writing:** Almost all students encounter some level of difficulty in adjusting to the specific demands of academic literacy needed at university level. Craig (1996), Mumba, Rollnick and White (2002), and Alfred, Dison and Hagemeyer (2000) identify this as the ‘gap’ between what is expected of students by lecturers and the general reality of what they can or do deliver. Students in the first year IS courses have difficulty with independent reading and research, display low levels of information literacy and find it difficult to select and analyse information appropriately (Johnston & Benvenuti, 2008). Writing is problematic for many students, and with the large numbers of students in the classes, writing tasks are difficult to manage as part of an assessment strategy.
- **Insufficient academic staff:** In common with first year IS courses at other universities in South Africa, the problem of a large number of students is often coupled with a shortage of academics (Benvenuti & Cohen, 2008). Several lecturers contributed to the teaching of the first year course, with all of them also responsible for teaching and coordinating courses in other years of undergraduate and postgraduate study.
- **Motivation differences between major and non-major students:** Students intending to major in IS also take this course as an introduction to the IS discipline and go on to undertake a second semester-long course in their first year of study before going on the IS courses at second and third year level. These students can be expected to have very different attitudes towards the course to their non-major peers and would need to be able to retain and build on the knowledge and skills taught in this course in subsequent years of study. Furthermore, as Biggs (2003) highlights, many of the students in our classes are not the traditional ‘academic’ university student looking for an education, but represent the new style student who is looking for a qualification and, ultimately, a job. This appears to be true of both the major and non-major students in the IS course.

- **Varied degrees of computer literacy:** Students embarking on this course are very varied in terms of their degree of computer literacy and prior knowledge and experience in working with technology. The group varied from students who had studied computer-related courses at school or university, students who owned their own computers and had grown up with technology in their schools and homes, and students with very little or, in some cases, absolutely no prior interactions with computers and technology.
- **Broad mix of knowledge types and skills:** The course outcomes cover a broad theoretical base of discipline-specific knowledge as well as focusing on developing procedural knowledge relating to the use of popular office productivity software packages. This requires a range of different assessment tasks.

These challenges are not dissimilar to those facing many first year introductory university courses, both locally in South Africa and elsewhere in the world.

## Evaluating existing course assessment practices

At the start of the research project the research team undertook an evaluation of the current research practices of the IS course. This involved an examination of the assessment tasks and course curriculum, together with interviews with the lecturers involved in teaching on the IS course.

The assessment tasks were found to consist largely of traditional timed assessments using multiple choice questions (MCQ), definitions and True/False assertions, with the final exam counting 50%, and three class tests together counting 24%. Practical assignments relating to the software package aspect of the course contributed the remaining 26% of the final overall mark. A detailed assessment of the exam questions using Anderson's (2005) revised Bloom's Taxonomy revealed that the majority of the exam made low cognitive demands on the students (Benvenuti & Cohen, 2008), while the assignments were also undemanding and required little more than the straightforward application of procedural knowledge in familiar contexts. A comparative review of first year examination papers from various other South African university IS departments revealed a very similar pattern (Benvenuti & Cohen, 2008).

The interviews with course lecturers revealed the following issues and concerns:

- Low pass rate.
- Low numbers continuing with major, i.e. continuing on to IS II.
- Large volume of 'boring' content is covered, resulting in no real depth and no real foundation taken through to second year.
- Great variance in prior knowledge.
- Lecturers feel limited by the volume of content and the large number of students.
- Students are passive and don't readily engage with material.
- Students seem largely driven by assessment requirements with effort based on perceived return. MCQs form a large part of exams, so students tend to 'cram facts' before exams.
- Students have a high workload so adopt strategic engagement with learning and assessment tasks.
- Material can be quite complex and 'alien', so students adopt rote learning without understanding, but are still able to pass exams and the course.

Based on the analysis of the assessment tasks and the issues raised by the academics in the interview sessions, a two-pronged approach to improving the overall assessment practices in the course was adopted. One aspect of the research project would focus on the ongoing assessment tasks undertaken by students during the course of the module, and is presented in some detail in this chapter. A separate, but related effort, focusing on final summative examinations was undertaken in parallel, and the early results of this work are discussed in chapter 2.

## Early interventions and evaluations

Using the interviews and preliminary data as a guide, and based on anecdotal evidence of lecturers, some workshops involving all the proposed IS I lecturers and the course coordinator were held, which resulted in a new set of assessment tasks being proposed for the following year.

The proposed tasks allowed for a wider variety of assessment tasks that would develop and assess a broader set of knowledge and skills, with certain tasks providing the opportunity for a more contextually embedded assessment. It was proposed that the new assessment tasks would be analysed in depth once implemented. Table 1.1 shows the composition of the new assessment tasks.

**Table 1.1: Composition of the new assessment tasks**

	Contribution to final mark	Task Description	Comment
Test and Exam	40 %	<i>See chapter 2 for details</i>	
InfoQuest	5 %	20 clues for information that must be found using either the Internet or other resources	Requires use of skills that form part of the course outcomes as well as covering some areas of content
Computer Literacy	10 %	Computer-based learning and assessment to cover the computer literacy aspects of the course	Needed in order to cope with large numbers of students with little exposure to computers
MS Access Assignments	10 %	Application of procedural knowledge	Retested under test conditions in Computer Literacy Test
MS Excel Assignments	15 %	Application of procedural knowledge	Retested under test conditions in Computer Literacy Test
Newspaper/IT Publication	20 %	Team-based assessment task to create an IT-focused publication that uses MS Word, Excel and PowerPoint, as appropriate. The publication should contain articles relevant to an IT audience and should be professional in terms of content and presentation	Designed to develop skills and attitudes (including writing, research, information synthesis, teamwork, creativity, professionalism, layout, formatting and referencing. Content coverage included new technology, research and development; career and employment – roles, skills, responsibilities, opportunities; general IS and IT exposure – knowledge and awareness

The various assessment tasks that were introduced were chosen to help develop and assess particular aspects of the course objectives. Some tasks were designed to develop the knowledge or skill required while others were more summative in nature and used the knowledge or skills developed in other tasks. For example, the class test was very similar in nature to the final course exam, thereby signalling to students what would be required of them in the final exam paper. Similarly, the InfoQuest task was designed to familiarise students with using electronic and other resources to obtain information for their articles in the IT Publication, while the computer literacy test and the assignments developed the procedural knowledge needed to format and produce the layout of the IT Publication.

The IT Publication was conceptualised as a means to encourage student engagement in the course, support the development of academic and information literacy, and allow for the development of more depth in discipline related knowledge. The publication was specified as a collaborative team project (consisting of nine to 12 members), in which each member of the team was required to independently write an article on any IT-related topic, which was then combined with the articles of the other team members for a single publication. The team was responsible for collating the articles, and designing the covers, advertisements and tables of contents. The team also had to conduct an IT-focused survey for inclusion in the magazine. A plan specifying the proposed articles and survey was submitted early on in the task (Johnston & Benvenuti, 2008).

The first implementation of the new set of assessment tasks was viewed very much as the first cycle in an action research context. Following the first year of implementation, the following observations were made:

- There were too many assessment tasks, which created a heavy marking load for staff and contributed to the heavy workload on students. Lecturers were struggling to mark the tasks in time to get feedback to students with the result that most tasks were purely summative, not developmental or formative.
- Some of the tasks were not always appropriate for achieving the desired assessment objectives (see more on this below).

Through discussions with the lecturers and examination of the various assessment tasks, the following decisions were made:

- The InfoQuest task was removed, with some of its objectives and tasks built into the practical laboratory tasks.
- The computer literacy tasks and office productivity software assignments were redesigned to build towards the types of knowledge and skills that students would need in order to successfully tackle the IT Publication Task. In this way, students build towards this competency through a structured approach ensuring competency and confidence with the tools before having to integrate them all in the major task.
- A tutorial portfolio was introduced to prepare students for tests and exams. This is discussed more fully in chapter 2.
- The IT Publication Task was redesigned, with a careful focus on the course objectives and outcomes, using the concept of ‘authenticity’ as the guiding principle.

The table on the following page shows the updated set of assessment tasks for the IS course.

**Table 1.2: The updated set of assessment tasks for the IS course**

Assessment Task	Contribution to final mark	Task Description	Comment
Test (15%) and Exam (30%)	45 %	<i>See chapter 2 for details</i>	
Tutorial Portfolio	10 %	Computer-based assessment that requires students to complete and submit tutorial work <i>prior</i> to their scheduled tutorial. This tutorial work is based on prior exam questions for that aspect of the course but allows students to reference their notes and textbooks	Students receive immediate feedback on their tutorial work. All students not achieving 80% or higher on each open-book tutorial are required to attend a tutorial session on that work. The tutorial sessions model the problem solving and knowledge construction required to answer each question
Practical Computer Lab Work Portfolio	25 %	This section now includes the computer literacy test and builds towards the application of procedural knowledge	Retested under test conditions in Computer Literacy Test and used in the production of the IT Publication Task
IT Publication	20 %	Team-based assessment task to create an IT-focused publication that uses MS Word, Excel and PowerPoint, as appropriate. The publication should contain articles relevant to an IT audience and should be professional in terms of content and presentation	Designed to develop skills and attitudes (including writing, research, information synthesis, teamwork, creativity, professionalism, layout, formatting and referencing). Content coverage included new technology, research and development; career and employment – roles, skills, responsibilities, opportunities; general IS and IT exposure – knowledge and awareness

## The successes and challenges of the IT Publication Task

The IT Publication task was hailed by the majority of the students as ‘the best part of the course’ and described as being ‘fun to do’, ‘interesting’, ‘stimulating’ and ‘allowed us to be creative and express ourselves’. Lecturers were delighted by the motivation of the students, their engagement with the task, and the overall quality of the publications from a design and presentation point of view. It was seen as an excellent assignment through which to develop, demonstrate, and assess skills in word-processing, spreadsheet and publication software, all part of the skills and literacy aspect of the course (Johnston & Benvenuti, 2008).

When examining the IT Publication Task against the original goals and outcomes, and taking into account how the students performed in the task, the following problems were identified:

- Problems with the design of the task: the task was not specific enough, with students taking shortcuts, and not engaging sufficiently with the writing and academic literacy aspects of the task. It was very clear from the majority of the articles received that students neither understood what

was required in writing an article, nor had many demonstrated any evidence of academic literacy. The articles in the main were rambling strings of facts, with little evidence of understanding of structure, argument, summary or conclusion. Most students had used only one source for their article, and these sources were often ill-chosen (Johnston & Benvenuti, 2008).

- Plagiarism: many of the articles were found to contain plagiarised content, with some being little more than ‘cut and paste’ exercises.
- Group work was problematic, with students either carrying a lot of extra work or doing very little, either by choice or in response to other members’ efforts or lack thereof.
- The task was found to be time consuming to mark and administer, especially with regards to helping students form or join teams, sorting out problems with submissions that were incomplete, etc. The marking of a single publication could take between 45 and 60 minutes, and criteria had not been set or communicated to the students or markers prior to assessment.

## Looking to the principles of authentic assessment for guidance

### What is authentic assessment?

Assessment is considered to be more authentic when it directly measures how well a student is able to perform tasks that are intellectually demanding and reflective of the real world in which such students will one day operate. This is in contrast to the traditional approach to assessment, in which indirect measures of student performance (such as answering questions in exams) are used to infer a student’s ability to perform real world tasks (Wiggins, 1990). The concept of authenticity was first used by Archbald and Newmann (1988) in conjunction with intellectual or academic achievement. They argue that for a student to demonstrate authentic academic achievement they need to be able to construct knowledge through ‘disciplined enquiry’ and engage with tasks that have ‘some meaning or value beyond achieving success at school’.

Wiggins (1989) introduced the term ‘authentic assessment’ in which he emphasises ‘the situational or contextual realism’ of the assessment tasks. He further asserts (1990) that authentic assessments ‘require students to be effective performers with acquired knowledge’, ‘present the student with the full array of tasks that mirror the priorities and challenges found in the best instructional activities’ and focus on the extent to which a student ‘can craft polished, thorough and justifiable answers, performances or products’.

Newman’s focus on the authenticity of performance achievement in his work on criteria for authentic intellectual achievement (the product), and Wiggins’ focus on the procedures and types of tasks more likely to elicit complex intellectual performance (the procedure) together suggest a strong relationship between authenticity and assessment. The relationship between them can be viewed in different ways. Cumming and Maxwell (1999) define the major difference in the discourse as being whether one is considering the ‘assessment of authentic achievement’ or the ‘authentic assessment of achievement’. They argue that, while these two positions have a different focus or emphasis, they are related and dependent on each other. It makes little sense to focus on authentic assessment without ensuring that the assessment tasks relate to the academic goals of the teaching and learning processes (Newmann & Archbald, 1992 cited in Cumming & Maxwell, 1999). Furthermore, it would be impossible to achieve authentic assessment without carefully considering what intellectual achievement must be demonstrated in the assessment tasks and how.

## The problems and challenges associated with authentic assessment

Two important objections to the use of authentic assessment need consideration. The first of these relates to the relatively high time demands that authentic assessment tasks make on lecturers. Well designed authentic tasks require substantial planning and preparation, and can result in products or performances that take time to mark. The use of detailed rubrics and group-based tasks, as discussed later, can help to alleviate this. Complex tasks are also more likely to be reusable, so time invested in planning and designing a task could result in tasks being reused for several years, with only minor changes and adjustments required (Montgomery, 2002).

The other major concern is task reliability. As Gipps (2009) points out, increased validity is often associated with decreased reliability in complex tasks. With a balanced assessment strategy that ensures a mix of various types of assessment tasks, and an overall consideration of validity, such concerns need not necessarily stand in the way of more authentic measures of achievement. Detailed assessment rubrics and task moderation by other academics in the field should help to achieve reliable judgements.

## Redefining the IT Publication Task using principles of authenticity

In redefining the IT Publication Task, we focused on clarifying what ‘real-life tasks, performances or challenges that mirror those faced by experts in the particular field’ (Wiggins 1989, cited in Montgomery, 2002) were important for our students to experience and develop. Through discussion it was decided that, while the tests and exams would continue to be used to assess important content knowledge, the IT Publication Task could be used to develop and assess a wider range of desired course outcomes. A well constructed and authentic task could require students to produce a complex product while exposing them to the challenges of a typical group-based working environment.

The IT Publication was designed to develop and assess achievement relating to both the product and the process of creating the product:

- Product – informative writing for a non-technical business audience on relevant IS issues, a small-scale IT related survey, a business publication created through the expert use of office automation products.
- Process – team work, project planning and management, writing, independent research.

More detailed guidelines were provided in terms of expected article length, style of writing and genre. A list of areas from which article topics could be selected was also provided. Students were also required to work with an editing partner, with each student required to read and provide feedback on a draft of their partner’s article. The intention of the partnership is to learn both from receiving and giving feedback on writing. See Appendix 1.1.

While students were expected to demonstrate achievement in some aspects of the task, other aspects were introduced as more developmental. The achievement goals included writing a well-constructed article on an IS-related issue, conducting a small-scale survey, and producing a professional document using appropriate technology. The developmental goals were more focused on the environment and contextual challenges that students could expect to contend with in the working world, such as teamwork, conflict management and negotiation, time management, organisational skills, as well as some areas that students would focus on later in their academic career. These included project planning and management (especially for intending IS majors), independent research and writing.

## Assessing the task

Described as the Editorial Board, the first submission for the IT Publication required students to put in place some plans for their publication including:

- the name of the publication;
- the names of the team members;
- a brief description of the planned mini survey questions;
- the titles and brief descriptions of the articles that would be included in the publication;
- a mini project plan with the budget required, financial contributions, proposed meeting dates, draft submission dates, editing plans and dates.

Once submitted, the coordinator of the assessment task met with each group to discuss their plans and to assist them with any problems. This was seen as an important aspect of the task, as it allowed the lecturer to engage the students in discussion around their proposed topics and provide some guidance in terms of relevance, and the style and approach that were expected.

When assessing a complex task, it is important to establish the criteria for achievement in the task. In this task, a mix of individual and group marks adds to the challenge. Rust, Price and O'Donovan (2003) report that students' learning and long-term uptake is greatly improved by developing their understanding of both assessment criteria and the assessment process. Detailed rubrics were constructed at the time of the design of the task in order to clearly define the requirements for the various elements of the task. Montgomery (2002) suggests that a detailed rubric is an appropriate tool for authentic assessment tasks as it enables a clear understanding by both students and lecturers of what is expected for each aspect of the complex task. It also offers detailed feedback to the students on their achievement, in what might otherwise be a purely summative task as it is submitted at the end of the course.

The use of well detailed and carefully constructed rubrics is also essential in mitigating some of the costs or challenges of authentic assessment. While time consuming to construct initially, such rubrics ensure that assessors are clear on the performance criteria of the task when marking, and inter-marker reliability is enhanced. Detailed feedback is given even when large numbers of students or groups of students are involved, and marking time is significantly reduced. With student numbers of over 900, and groups of between 10 and 12 students, approximately 80 publications with a total of over 900 articles had to be marked in the period of two weeks. Appendix 1.2 and 1.3 contain the marking rubrics for the Editorial Board and publication respectively.

Process-related aspects of the publication were not specifically assessed. Instead, a mandatory reflective questionnaire was completed by each student and submitted with the publication. The questionnaire did not contribute any marks to the assessment task. Students were asked to comment on issues relating to working with an editing partner, managing the project and general comments relating to the task and experience. The student comments were used by the lecturers to refine the task and task briefing sessions, and intended to help students when approaching similar tasks in future courses.

## Reflecting on the IT Publication Task

The IT Publication Task has evolved over the past three years into a task that now offers students a more authentic means by which to develop and demonstrate complex achievement. While the lecturers are still fine tuning the task, it represents a major step forward from the narrow approach to assessment of a few years ago.

Students find the task challenging, but fun and rewarding, though many oppose having to work in teams. The majority of students identify the IT Publication as being the best part of the course, despite the difficulties and challenges they encounter along the way. Producing glossy, high quality magazines that look and feel professional represents a major and tangible achievement for all students, particularly for those who started the course with little or no prior computer experience.

With team-based work environments and collaborative work high on employer agendas, and the very real issue of class size, the team work aspect of the project will definitely not go away. Students move into very complex software analysis, design and development projects in their second year of study, and, in common with the working world, these are team-based too. The challenge for lecturers is to find ways in which to help students adjust to the team work situation and manage the problems and difficulties they encounter.

From the lecturers' point of view, the task is fulfilling many of the learning objectives of the course and is providing a far more satisfactory and valid demonstration of student learning and achievement. A continued focus for the improvement of this task will be on how to help students develop their writing skills using an editing partner, and on developing students' critical reading and thinking skills in order to better select and synthesise information for their writing.

## References

- Alfred, L, Dison, L & Hagemeyer, L, 2000. Developing a Sociology Foundation Course: An Analysis of Principles, Methods, and Materials. *South African Journal of Higher Education*, 14(3), 112–19.
- Anderson, WL, 2005. Objectives, Evaluation and the Improvement of Education. *Studies in Educational Evaluation*, 31, 102–13.
- Archbald, DA & Newmann, FM, 1988. *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School*. Reston, VA: National Association of Secondary School Principals.
- Benvenuti, S & Cohen, JF, 2008. An In-Depth View of MCQ-based Testing in Information Systems. *South African Computer Journal*, 42, 21–28.
- Biggs, J, 2003. *Teaching for Quality Learning at University* (2nd ed.). The Society for Research into Higher Education & Open University Press.
- Boud, D, Cohen, R & Sampson, J, 1999. Peer Learning and Assessment. *Assessment and Evaluation in Higher Education*, 24(4), 413–26.
- Brown, S & Knight, P, 1994. *Assessing Learners in Higher Education*. London: Kogan Page.
- Brown, G, Bull, J & Pendlebury, M, 1997. *Assessing Student Learning in Higher Education*. London: Routledge.
- Craig, A, 1996. Education for All. *South African Journal of Higher Education*, 10(2), 47–55.
- Cumming, JJ & Maxwell, GS, 1999. Contextualising Authentic Assessment. *Assessment in Education: Principles, Policies and Practices*, 6(2), 177–94.
- Entwistle, N & Entwistle, A, 1997. Revision and the Experience of Understanding. In F. Marton, D. Hounsell & N. Entwistle (Eds). *The Experience of Learning*. Edinburgh: Scottish Universities Press.
- Gibbs, G, 1999. Using Assessment Strategically to Change the Way Students Learn. In A. Brown & S. Glassner (Eds). *Assessment Matters in Higher Education*. Buckingham: Society for Research into Higher Education and Open University Press.
- Gibbs, G & Jenkins, A, 1992. An Introduction: The Context of Changes in Class Size. In G. Gibbs & A. Jenkins (Eds). *Teaching Large Classes in Higher Education: How to Maintain Quality with Reduced Resources*, 11–22. London: Kogan Page.
- Herrington, J & Herrington, A, 1998. Authentic Assessment and Multimedia: How University Students Respond to a Model of Authentic Assessment. *Higher Education Research and*

*Development*, 17(3), 305–22.

Johnston, KA & Benvenuti, S, 2008. News You Can Use. Proceedings of the Informing Science & IT Education Conference (InSITE) 2008.

Montgomery, K, 2002. Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, 50(1), 34–9.

Mumba, FK, Rollnick, M & White, M, 2002. How Wide is the Gap Between High School and First-year Chemistry at the University of the Witwatersrand? *South African Journal of Higher Education*, 16(3), 148–56.

Newmann, F, Brandt, R & Wiggins, G, 1998. An Exchange of Views on Semantics, Psychometrics, and Assessment Reform: A Close Look at ‘Authentic Assessments’. *Educational Researcher*, 27(6): 19–22.

Rowntree, D, 1987. *Assessing Students: How Shall we Know Them?* London: Kogan Page.

Rust, C, Price, M & O’Donovan, B, 2003. Improving Students’ Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education*, 28(2), 147–64.

Terwilliger, J, 1997. Semantics, Psychometrics, and Assessment Reform: A Close Look at ‘Authentic Assessments’. *Educational Researcher*, 26(8), 24–7.

Wiggins, G, 1989. A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappan*, 70, 703–13.

Wiggins, G, 1990. *The Case for Authentic Assessment*. ERIC Clearinghouse on Tests, Measurement and Evaluation, Washington, DC. (ED 328 611)

# Appendices

## Appendix 1.1 – IT Publication student brief

### **Fundamentals of information systems – IT publication**

You are required to produce an IT Publication (magazine) in groups of 10 to 12 students (no more and no less!)

The learning objectives are to:

- develop research skills
- develop writing skills
- develop the skill of synthesising information
- develop team work and project management skills
- demonstrate creativity
- stimulate interest in new and exciting technology
- expose you to career and employment opportunities
- broaden your general knowledge of IS and IT

As well as:

- to get good marks
- and last, but not least, to have fun!

The publication consists of 2 deliverables:

- an Editorial Board – Due Date 17 March, 17h00
- the final magazine – Due Date 5 May, 17h00

Both deliverables should be posted into the Information Systems IA wooden box, outside the secretary's office on the 1<sup>st</sup> floor of the Commerce, Law & Management building. If your magazine does not fit into the box, you may hand it to the secretary. There will be bonus marks for early submissions.

Groups must see Linda Spark (Room 126, 1<sup>st</sup> floor, Commerce, Law & Management building) during the week 31 March to 4 April to receive feedback about their Editorial Board. Where possible, the entire group should attend.

The **Editorial Board** should include the following:

- a completed Editorial Board Assessment form. This form must be completed on WebCT, printed off and signed by each group member and stapled as the front cover of your Editorial Board submission
- the detail and questions of a survey that you will carry out (see Notes below – point 1)
- the articles that will be covered in your publication (see Notes below – point 2)
- a mini project plan (see Notes below – point 3)

The Editorial Board must meet the following specifications:

- typed on A4 sheets (portrait)
- stapled in the top left corner (do not submit in a plastic sleeve or any type of binder or file)
- properly proofread for typing and spelling errors

The **final magazine** must meet the following specifications:

- typed on A4 sheets (portrait) and bound
- properly proofread for typing and spelling errors
- include the results of a survey (see Notes below – point 1)
- contain your articles – the number of these should be the same number as the number of people in your group and should be written by the person specified in the Editorial Board (see Notes below – point 2)
- include advertising (see Notes below – point 4)
- include page numbers and a table of contents
- no more than 10% of the content of any article may be directly copied ('cut-and-paste') from any other source (exclude advertising from this percentage) and all sources of information must be referenced (see Notes below – point 5)
- attach an individual response to a reflective questionnaire (see Notes below – point 6) – do not bind within the magazine; either slip it in or use a paperclip
- you may include additional articles of interest (bonus features) such as crosswords, cartoons, IT movie reviews, IT book reviews, IT related classifieds, letters pages, etc. Bonus features do not count as articles in the magazine, but will earn bonus marks
- you may choose any fonts, colours and sizes, but the magazine should be clear to read
- attach the following forms: (do not bind within the magazine; either slip in or use a paperclip)
  - a Summary Assessment form – this form must be completed on WebCT, printed off and signed by each group member and handed in with your final magazine
  - an overall publication rubric (marking memo)
  - an individual contribution form for each member – these should be completed on WebCT

## Notes

### 1. Survey

This must be carried out by **each** member of the group on **any IT topic or issue**, on a minimum of 5 students on campus (excluding members of your group) and the results of the survey must be analysed/discussed/presented in the final publication. There should be at least 5 questions in your survey. You should use graphs or charts to present your results.

The Editorial Board must include a short background of about 3 paragraphs about the IT topic/issue you have chosen to survey together with the questions you will ask.

The final magazine should again present the background to your survey, present the results, some analysis and discussion of these and some recommendations.

### 2. Articles

For the Editorial Board, you must list who will be responsible for each article that your magazine will be covering. Give the title of the article, 2 to 3 lines explaining what this article will be about and the author's editing partner (your editing partner is the person who will be responsible for helping you ensure that your article meets the requirements)

In the final magazine, each article should have an introduction, body and conclusion. The article should be 250 to 300 words in length. The name of the person who wrote the article should appear at the top of the article. Your articles must be chosen from the following topics:

- case study of an organisation that has achieved breakthroughs in performance from IS/ improved customer service/improved managerial decision making/achieved advantage
- advice to small businesses on how IT can help them do some of the above
- review and comparison between two PCs in the same price range

- review and comparison of two software applications
- review and comparison between two database management systems
- article on the history and growth of the Internet and projections for the future
- article on the latest web-based applications/uses of the web (business or social)
- who's who on the Web
- ERP II
- review of business intelligence
- use of MIS or DSS in a specific type of business, e.g., banking, retail, education
- the role of the business user in improving systems development success
- the digital divide
- recent hacking incident
- security threats and what to do about them (advice for small businesses)
- an IT career review
- a profile of an IT vendor
- IT business trends
- the relevance of ethics to IT professionals

### **3. Mini Project Plan**

This plan should set out the following:

- the budget required
- each person's financial contribution
- proposed meeting dates
- draft submission dates
- editing plans and dates
- final meeting date for the entire group to sign off and submit the magazine

### **4. Advertising**

This does not count as a section of the magazine. Adverts can be cut out from other papers or magazines or be produced by your group (the adverts must be for IT products or services, but they do not have to be real products or services). Advertising should not take more than 20% of your publication.

### **5. References**

These must show where you got your information by stating the author, the website address or name of book/journal/magazine/newspaper and the date of the book/journal/ magazine/ newspaper or the date the website was accessed, for each article in the publication. These should appear at the bottom of the article. You are expected to refer to at least 2 sources of information.

### **6. Reflective Questionnaire**

Each member of the group should complete this questionnaire on WebCT (individually), print it off and attach (not bound) to the magazine. There will be a penalty if these forms are omitted.

## Appendix 1.2 – Editorial Guide assessment criteria

Editorial Board (25% of final mark)

Total \_\_\_\_ / 50

Professionalism and Organisation

Max Mark 30

Mark Awarded \_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5)	(5)	(3-4)	(1-2)	(0)	
Cover page with student names, numbers and publication title	All requirements met	Most requirements met	Several requirements not met	Requirements not met	
(10)	(9-10)	(7-8)	(5-6)	(0-4)	
Project Plan	Well planned and laid out All requirements met: • Budget and contributions • Meeting dates and plans • Editing dates and plans • Signoff Meeting date	Good planning effort and layout Most requirements met • Budget and contributions • Meeting dates and plans • Editing dates and plans • Signoff Meeting date	Plan missing some important aspects • Budget and contributions • Meeting dates and plans • Editing dates and plans • Signoff Meeting date	Planning not adequately done • Budget and contributions • Meeting dates and plans • Editing dates and plans • Signoff Meeting date	
(15)	(12-15)	(8-11)	(5-7)	(0-4)	
Proposed Articles	Majority of articles properly presented • Student name • Title • Brief explanation	Several articles not properly presented • Student name • Title • Brief explanation	Many articles not properly presented • Student name • Title • Brief explanation	Overall poor presentation of proposed articles • Student name • Title • Brief explanation	

Survey

Max Mark 20

Mark Awarded \_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(10)	(9-10)	(7-8)	(5-6)	(0-4)	
Background Explanation	Both the purpose and proposed implementation of the survey are well explained	Both the purpose and proposed implementation of the survey are adequately explained	The explanation of either the purpose and/or proposed implementation of the survey are inadequately explained or missing	The explanation of both the purpose and implementation of the survey are missing or inadequate	
(10)	(9-10)	(7-8)	(5-6)	(0-4)	
Suitability and design of questionnaire	Easy to use/answer the questions; questions provide detailed information relating to the survey issue	Easy to use/answer the questions; questions provide sufficient information relating to the survey issue	Questions provide sufficient information relating to the survey issue	Questions provide insufficient information relating to the survey issue	

## Appendix 1.3 – IT Publication assessment criteria

**Overall Publication (25% of final mark) Sub Total**

### Professionalism and Organisation

**Max Mark 20**

**Mark Awarded \_\_\_\_**

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(10) Table of contents & page numbering	(9-10) Page numbers Detailed table of contents with reference to page numbers	(7-8) Page numbers Detailed Table of contents but no reference to page numbers OR Limited table of contents with reference to page numbers	(5-6) Page numbers Table of contents Lack of detail in ToC AND/OR no reference to page numbers	(0-4) No Page numbers AND/OR No table of contents	
(10) Proofread – grammar & spelling	(9-10) Few if any spelling or grammar mistakes	(7-8) A few minor grammar or spelling mistakes	(5-6) Contains some spelling or grammar mistakes	(0-4) Contains many spelling AND/OR grammar mistakes	
Penalty (-10) Meets submission criteria: format, on time, meets all requirements	(0) Submitted on time Correct format Assessment Forms completed	(0 to -2) Minor format problems OR Missing requirements	(-3 to -5) Either submitted late OR Incorrect format OR Missing requirements	(-6 to -10) Submitted late AND Incorrect Format OR Missing Requirements	

### Content, Creativity and Presentation

**Max Mark 20**

**Mark Awarded \_\_\_\_**

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Content	(5) All content is relevant, up-to-date, varied, and of interest to an IT audience	(4) Most of the content is relevant, up-to-date, varied, and of interest to an IT audience	(2-3) Less than half the content is relevant, up-to-date, varied, and of interest to an IT audience	(0-1) The majority of the content is either not relevant, up-to-date or of interest to an IT audience	
(5) Cover	(5) Has many of the following characteristics: Exciting, well laid out, appeals to an IT audience, highlights interesting articles, title & design IT focused	(4) Has some of the following characteristics: Exciting, well laid out, appeals to an IT audience, highlights interesting articles, title & design IT focused	(2-3) Has a few of the following characteristics: layout is not professional, is difficult to read, does not provide information on content, title and design are not IT focused	(0-1) Has several of the following characteristics: layout is not professional, is difficult to read, does not provide information on content, title and design are not IT focused	

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Layout	(5) The layout is professional, appealing to the eye, easy to read, articles attract interest, use of headings, bylines, etc.	(3-4) The layout is generally professional, appealing to the eye, easy to read; some use made of headings, bylines, etc.	(2) The layout is not professional – more like an essay than a publication OR difficult to read OR lack of headings, etc.	(0-1) The layout is not professional – more like an essay than a publication AND/OR difficult to read AND/OR lack of headings, etc.	
(5) Consistency and flow	(5) The publication is consistent in layout and style throughout, flows well in terms of organization and order of articles	(3-4) The publication is fairly consistent in layout and style, generally flows in terms of organization and order of articles	(2) The publication lacks consistency in layout and style OR lacks flow in terms of organization and order of articles	(0-1) The publication lacks consistency in layout and style AND lacks flow in terms of organization and order of articles	

#### Adverts + Bonus Features

Max Mark 20 (includes 10 bonus marks)

Mark Awarded \_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Appropriate choice of adverts	(5) Adverts are all related to IT and particularly to articles in this publication	(3-4) Adverts are all related to IT	(1-2) Most adverts are related to IT	(0) Adverts not related to IT OR no adverts included	
(10) Bonus features	(8-10) Publication includes bonus features that are interesting AND/OR relevant to your publication audience	(5-7) Publication includes bonus features with some interest AND/OR relevance to your publication audience	(1-4) Publication includes a few bonus features OR includes bonus features that are not of particular interest or relevance to your readers	(0) No bonus features included	
(3) Integration into publication	(3) Adverts and/or bonus features well positioned relating to articles with some on pages with related articles	(2) Adverts AND/OR bonus features well positioned relating to articles OR Adverts AND/OR bonus features on pages with related articles	(1) Adverts randomly placed in publication	(0)	
(2) Appropriate % of adverts/bonus features	(2) 10–20% of publication consists of adverts/bonus features		(1) >10% of publication consists of adverts/bonus features	(0) No adverts/bonus features	

## Survey (25% of final mark)

Max Mark 50

Mark Awarded \_\_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Relevance and usefulness	(5) Relevant to IT and particularly to articles in this publication; provides useful information	(3-4) Relevant to IT and particularly to articles in this publication or provides useful information	(1-2) Relevant to IT	(0) Not relevant to IT	
(10) Background explanation	(5) Both the purpose and implementation of the survey are well explained	(4) Both the purpose and implementation of the survey are adequately explained	(2-3) The explanation of either the purpose and/or implementation of the survey are inadequately explained or missing	(0-1) The explanation of both the purpose and implementation of the survey are missing or inadequate	
(5) Suitability and design of questionnaire	(5) Easy to use/answer the questions; questions provide detailed information relating to the survey issue	(3-4) Easy to use/answer the questions; questions provide sufficient information relating to the survey issue	(2) Questions provide sufficient information relating to the survey issue	(0-1) Questions provide insufficient information relating to the survey issue	
(20) Analysis and discussion of results	(5) The results of the survey are discussed, analysed and summarised	(3-4) The results of the survey are discussed AND/OR analysed AND/OR summarised	(2) Some discussion or analysis or summarisation of results	(0-1) Little or no discussion, analysis or summarisation of results	
(5) Presentation of results	(5) Survey results well presented using graphs AND/OR charts – easy to understand or self explanatory	(3-4) Survey results presented using graphs AND/OR charts – mostly easy to understand or self explanatory	(2) Results difficult to interpret – little or no use made of graphs, charts or summaries	(0-1) Results poorly presented – no use of graphs, charts or summaries	
(5) Recommendations	(5) Good recommendations based on most of the findings are made	(3-4) Recommendations based on several of the findings are made	(1-2) Recommendations based on some of the findings are made	(0) No recommendations are made	

## Individual Contribution (25% of final mark)

### Professionalism, Organisation and Presentation

Max Mark 10

Mark Awarded \_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Proofread – grammar & spelling	(5) Few if any minor spelling or grammar mistakes	(3-4) A few minor grammar or spelling mistakes	(1-2) Contains some spelling or grammar mistakes	(0) Contains many spelling AND/OR grammar mistakes	
(5) Layout Readability Creativity	(5) Well laid out with excellent use of headings, subheadings, quotes, bylines; excellent use made of paragraphs, fonts and formatting – easy to read; Creatively designed – excellent use made of appropriate artwork or graphics	(3-4) Well laid out with some good use of headings, subheadings, quotes, bylines; Good use made of paragraphs, fonts and formatting – fairly easy to read; Creatively designed – some good use made of appropriate artwork or graphics	(1-2) Fairly well laid out, some use made of headings, subheadings, etc. Some use made of paragraphs, fonts and formatting Limited use of artwork or graphics OR artwork and graphics not really relevant to article	(0) Article not professionally laid out, limited use made of headings, subheadings, etc. Minimal use made of paragraphs, fonts or formatting – difficult to read. No use made of artwork or graphics OR artwork and graphics not relevant to article	

### Content

Max Mark 40

Mark Awarded \_\_\_\_

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) Interest and relevance to IT	(5) The article is both very relevant and of interest to an IT audience	(3-4) The article is relevant and of interest to an IT audience	(1-2) The article has some relevance or interest to an IT audience	(0) The article is of little or no real interest or relevance to an IT audience	
(25) Content, structure and flow of article Genre	(25-30) The article is well structured with an introduction and conclusion, and a logical flow and development of ideas Appropriate genre or style used	(18-24) The article has fair structure and a logical flow of ideas Awareness of appropriate genre or style demonstrated	(12-17) The article lacks structure AND/OR flow of ideas Appropriate genre or style not demonstrated	(0-11) The article is unstructured with little flow or development of ideas Lack of awareness of appropriate genre or style	
(5) Originality of work	(5) The article contains the student's own thoughts, ideas and opinions relating to the facts presented	(3-4) The article contains some evidence of the student's own thoughts, ideas and opinions	(1-2) There is little evidence of the student's own thoughts, views or ideas in the article	(0) There is no evidence of the student's own thoughts, views or ideas in the article	

	Excellent	Very Good	Acceptable	Needs additional work	Mark
(5) References	(5) Information is accurate and relevant to the article. Several good sources have been used (>2) References properly presented	(3-4) The majority of information is accurate and relevant to the article. Sufficient sources have been used(2) References properly presented	(1-2) Insufficient or poor quality sources have been used (<2) AND/OR references incorrectly presented	(0) Insufficient AND/OR poor quality sources have been used (<2) References incorrectly presented	

# Using MCQ-based assessment to achieve validity, reliability and manageability in introductory level large class assessment

*Susan Benvenuti*

## Introduction

Growing student numbers and the related increased workload for academic staff has resulted in assessment practices that are not always ideal. Assessment tasks seem often to be designed with more cognisance of marker workload than sound assessment principles, resulting in the frequent reliance on multiple choice question (MCQ) based assessment in large introductory courses.

In this chapter, we look at another aspect of the broader research project that examined issues relating to assessment in an introductory course in Information Systems (IS) as discussed in chapter 1. In particular, we adopt a pragmatic stance that argues that the use of MCQs as a strategic response to assessment of very large introductory level courses can still deliver valid and reliable judgement of student learning and understanding, while at the same time allowing for manageability of final assessment given resource and time constraints. With careful attention to question design, this type of assessment is able to test various types of knowledge (factual, conceptual and procedural) at different levels of cognitive demand (recall, application, etc.) and is able to provide a valid and reliable judgement of student learning.

## Background and context

When the initial research project began, the first year IS course was selected as an appropriate, representative course for the Commerce, Law and Management (CLM) Faculty as it fulfilled the criteria of being offered at first year level, having a large number of students, and being a compulsory course for all intending BCom students. At the start of the project, the class size was approximately 250 and was expected to grow to over 400 the following year; a year later the numbers swelled to over 900.

This growth is similar to that faced by the majority of both regional and international universities, and is coupled with a shortage of academic staff in many disciplines, including IS (Benvenuti & Cohen, 2008). The resulting higher student-lecturer ratio has done little to encourage good assessment practices, with many academics relying more and more heavily on assessment tasks that are manageable in terms of marking loads.

As discussed in chapter 1, during the initial stage of the research project the assessment tasks for the course were found to be very homogeneous in nature, with 74% of a student's final course

mark being determined by tests and exams. Furthermore, these assessment tasks consisted largely of MCQs, incomplete definitions and True/False assertions. A detailed analysis of the questions from the previous year's exam using Anderson's (2005) revised Bloom's Taxonomy revealed that the majority of the questions were definition based, and involved factual recall. A comparative review of questions taken from examination papers from similar courses at other South African universities showed little difference (Benvenuti & Cohen, 2008).

This chapter presents a brief discussion of the results of the mini research project and focuses on how lecturers on the course now work with MCQs as part of their overall teaching, learning and assessment strategy. A more detailed discussion of this mini research project can be found in Benvenuti and Cohen (2008).

## The research project

Following on from the early analysis of the assessment practices for the IS course, a complete redesign of the assessment strategy for the introductory IS course was undertaken as described in chapter 1, resulting in traditional timed tests and exams now counting only 45% of the course mark. In addition, a mini research project was undertaken in order to investigate whether MCQ testing could provide a valid, reliable and pragmatic response to the realities of growing student numbers and the constraints imposed by resource shortages.

The mini research project asked the following questions:

**Research Question 1** – To what extent are MCQs being used in the assessment of higher level cognitive engagement with knowledge in the foundation information systems courses of South African tertiary institutions?

**Research Question 2** – Is it possible to design MCQs that assess students at deeper levels of cognitive ability?

**Research Question 3** – Is it possible for MCQs to distinguish between those students who have constructed conceptual relationships between concepts in the field (high level cognitive ability) and those who have only been capable of memorising isolated facts (low level cognitive ability)?

In other words, we were hoping to find ways in which to 'capitalise' on the benefits of using MCQs (reliability, test automation, detailed and very quick feedback, etc.) while overcoming issues such as low validity and lack of discrimination between performances of poor, average and good students (Benvenuti & Cohen, 2008).

## Research Question 1

*To what extent are MCQs being used in the assessment of higher level cognitive engagement with knowledge in the foundation information systems courses of South African tertiary institutions?*

In order to answer the first research question, a request was sent to South African Computer Lecturers Association (SACLA) members to share their past papers with the researchers for the purposes of this analysis. Five institutions responded and a total of 231 unique MCQs were classified using Anderson's (2005) revised Bloom's Taxonomy, as shown in Table 2.1. The examination papers, from which the questions were drawn, were all used in the summative assessment of students in the foundation (first year) information systems/informatics courses, and all employed the use of multiple choice questions (Benvenuti & Cohen, 2008).

**Table 2.1: Classification of MCQs taken from first year IS exam papers from several South African universities using Bloom's Revised Taxonomy (Benvenuti & Cohen, 2008)**

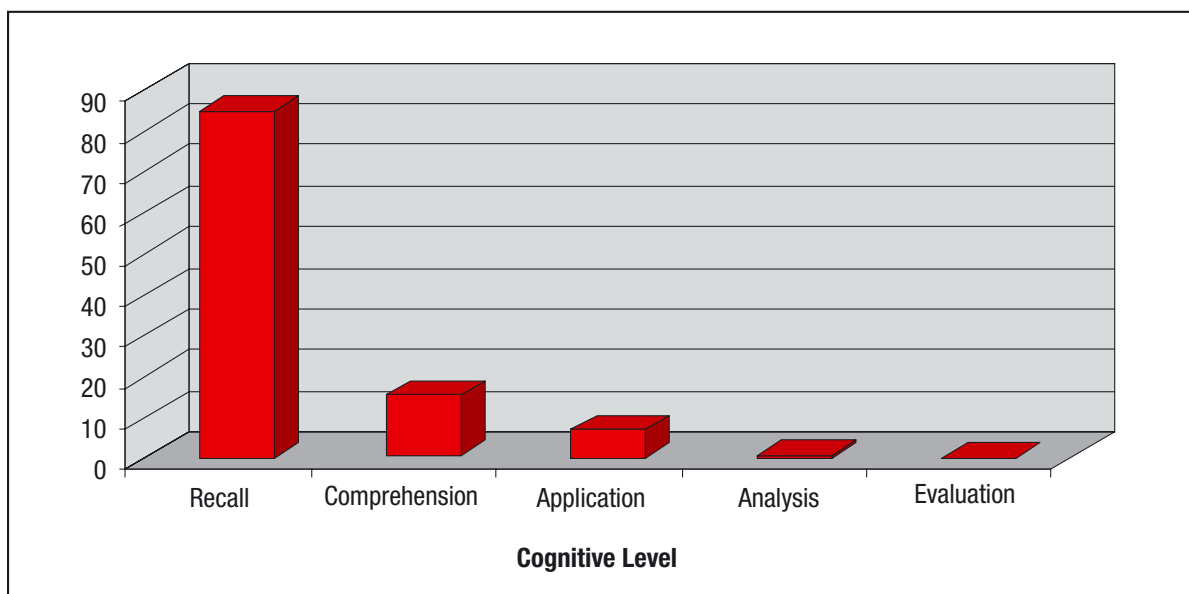
The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember/ Recall	2. Understand	3. Apply	4. Analyse	5. Evaluate	6. Create
Factual Knowledge	171	9				
Conceptual Knowledge	16	19	5	2*		
Procedural Knowledge	1		8			
Metacognitive Knowledge						

\* These questions were from a second-year level course that was included in the pool of examination papers. This was due to the role it played as an introductory/foundation IS course.

The analysis revealed that 74% of the questions demanded little more than the recall of factual knowledge presented in a typical introductory textbook, with 81% in total at the cognitive recall level (see Figure 2.1). Only 12% of the questions required students to demonstrate comprehension of factual or conceptual knowledge. This was generally achieved by requiring students to interpret data/facts or correctly identify the use of a concept in an example not previously seen (Benvenuti & Cohen, 2008). Woodford and Bancroft (2005) define such questions as requiring students to have understood what they have learned and translate that knowledge into a new context.

Just over 6% of questions tapped into the application and analysis levels, which should require students to solve problems by applying knowledge, facts, techniques and rules (application) or examining information, breaking it into parts, identifying patterns, causes, relationships, analysing effects and making inferences (analysis) (Woodford & Bancroft, 2005). It was also found that the majority of application-level MCQs involved the application of programming/software application rules or formulae, as opposed to the application of management information systems principles.

**Figure 2.1: Percentage of MCQs addressing different levels of cognitive skills across all the exam questions sampled (Benvenuti & Cohen, 2008)**



## Looking to the assessment literature for guidance

The bleak results of our analysis of the MCQ assessment tasks confirmed that the cognitive demands made on our students were lower than desired, and that students were neither encouraged to develop sufficient depth in learning or able to demonstrate it should it exist (Benvenuti & Cohen, 2008).

As we are frequently reminded,

From our students' point of view, assessment always defines the actual curriculum. In the last analysis, that is where the content lies for them, not in lists of topics or objectives. Assessment sends messages about the standard and amount of work required, and what aspects of the syllabus are most important (Ramsden, 1992:187).

Biggs (2003) comments that the exclusive use of MCQs 'greatly misleads' students in terms of 'the nature of knowledge', with all knowledge, regardless of whether it be detailed facts or overarching concepts or principles, reduced to the same value or mark. He cites Lohman (1993:19):

There is no need to separate main ideas from details; all are worth one point. And there is no need to assemble these ideas into a coherent summary or to integrate them with anything else because that is not required.

All the questions analysed in our study, regardless of cognitive demand, were equally valued as a demonstration of knowledge (assigned one mark) by the assessors. This not only fails to consider the time requirements needed by students to answer higher level questions, but, more importantly, signals to students that the factual recall of small detail is as important as broader understanding of concepts, ability to relate concepts to one another, and ability to apply those in given problems and scenarios.

'Assessment therefore tends to determine what the student learns ("is this in the exam?"), what the student does ("does this count for marks?"), how the content is learned or covered (memorisation of facts versus attempts to construct conceptual frameworks with relationships, etc.) and the degree to which the knowledge (or set of isolated facts) is retained and usable by students (application, generalisation and further development)' (Benvenuti & Cohen, 2008:22). Assessment tasks and criteria can and do signal what is valued by assessors, or, as Rowntree (1987:31) suggests, 'what the system requires students to *do* to survive and prosper'. We need to ensure that we are sending the right signals.

## Using MCQs as a strategy for assessing large classes

Assessment is recognised as playing many roles, including judgement of mastery or achievement, and providing feedback to teachers or students on progress or understanding. While the evidence and literature in the above discussion suggest that MCQ testing does in general perform this role adequately, central too, to discussions on assessment practices and strategies, is careful consideration of validity, reliability and manageability.

Validity considers the degree to which an assessment task or question is a valid means of judging the achievement of a required learning outcome or providing feedback on knowledge or progress. If, for example, it is important for a student to be able to write a computer program to perform certain calculations, it would not be sufficient for students to demonstrate merely that they know how to do the calculation on paper.

Reliability looks at the degree to which we could confidently predict the same result or judgement based on the sample provided by the assessment tasks or questions. Would a student either passing or failing the task be likely to have achieved the same result on another task testing the same ability or knowledge, or if the assessment had been conducted by another assessor?

Manageability of assessment tasks or strategies considers issues such as how usable and affordable the assessment tasks are in terms of the feedback they provide, and the availability of resources, including time, manpower and facilities (Gibbs & Jenkins, 1992; Knight, 2001).

Given that one of the major issues driving this research project (and shared by all the universities considered in the comparative analysis in this research), is that of large classes and tight resources, manageability of assessment should be given as strong a focus as validity and reliability when considering assessment strategies.

Strategies for large class assessment include reducing the assessment load and the delegation of marking (Brown, 2001). These strategies often employ ICTs to produce assessment tasks that allow for relatively easy marking while maintaining reliability. MCQs are frequently used in courses with large numbers, broad curricula and resource constraints, as a strategic means by which to assess. Our research and investigations alone revealed that many South African universities use MCQs to test the majority of the course content in final summative examinations in large first year courses across all faculties.

MCQs have many positive aspects, including high reliability, high manageability in terms of ease of marking (through delegation or automation), ease of use in implementation (automation and availability of questions from textbook test banks), as well as immediate and detailed feedback to lecturers and students (Benvenuti & Cohen, 2008:23).

On the negative side, both our research and the literature suggest that most MCQs appear to be set at low levels of cognitive demand and they may be time consuming to set if well constructed and unambiguous.

## Research questions 2 and 3

2. *Is it possible to design MCQs that assess students at deeper levels of cognitive ability?*
3. *Is it possible for MCQs to distinguish between those students who have constructed conceptual relationships between concepts in the field (high level cognitive ability) and those who have only been capable of memorising isolated facts (low level cognitive ability)?*

In order to examine these questions, a small study was undertaken in which two tutorial groups of students (119 students in total) were asked to answer a set of 20 MCQs on the topic of the Internet and the World Wide Web (www). The students completed the MCQs in their own time, and were allowed access to textbooks and class notes.

The Internet and www was a topic area in which relatively few questions reviewed in phase 1 went beyond the level of knowledge recall. This is an important topic in which the goal is to expose students to sufficient content in order for them to appreciate and understand the implications of the Internet and www for individuals and society, their application for business, and issues surrounding their use. The topic should demystify the Internet and provide students with the foundation knowledge needed for future topics, including e-commerce, web programming and network management. Students hold several typical misconceptions in this area, including the relationship between the Internet and www, as well as the relationship between IP addresses and URLs. Weaker students 'struggle to link all the pieces of the Internet puzzle – it is an area in which even computer literate students will encounter unfamiliar terminology (technical jargon) such as routers, packets, TCP/IP, and client/server architecture' (Benvenuti & Cohen, 2008:24).

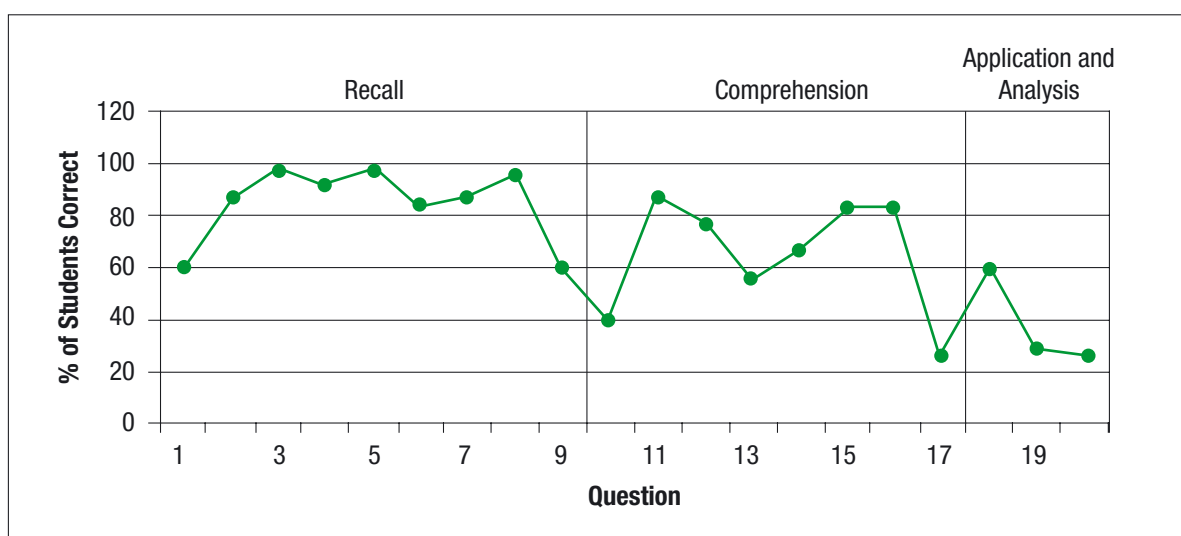
Prior to the study, lecturers attended a workshop on the use of MCQs and the setting of higher order MCQs based on the experiences of other disciplines, such as education and psychology. Thereafter, lecturers set a mix of questions for the topic area, including several that focused on drawing out misconceptions, as well as testing for student understanding of the big picture, as opposed to simply testing knowledge of isolated facts.

Most students were expected to do well in the sections requiring simple recall due to the open-book nature of the exercise. Only those students who adequately understood concepts were expected to perform well in questions testing comprehension. Furthermore, comprehension would be a necessary prerequisite for performance on questions requiring application and analysis.

## The results of the study

As expected, the class on average scored very well on recall questions, with students performing less well on questions requiring comprehension, as seen in Figure 2.2. Only 60% of students were able to answer the application question correctly, and less than 30% of students were able to answer the analysis questions correctly. For more detailed analysis of the results, see Benvenuti & Cohen (2008).

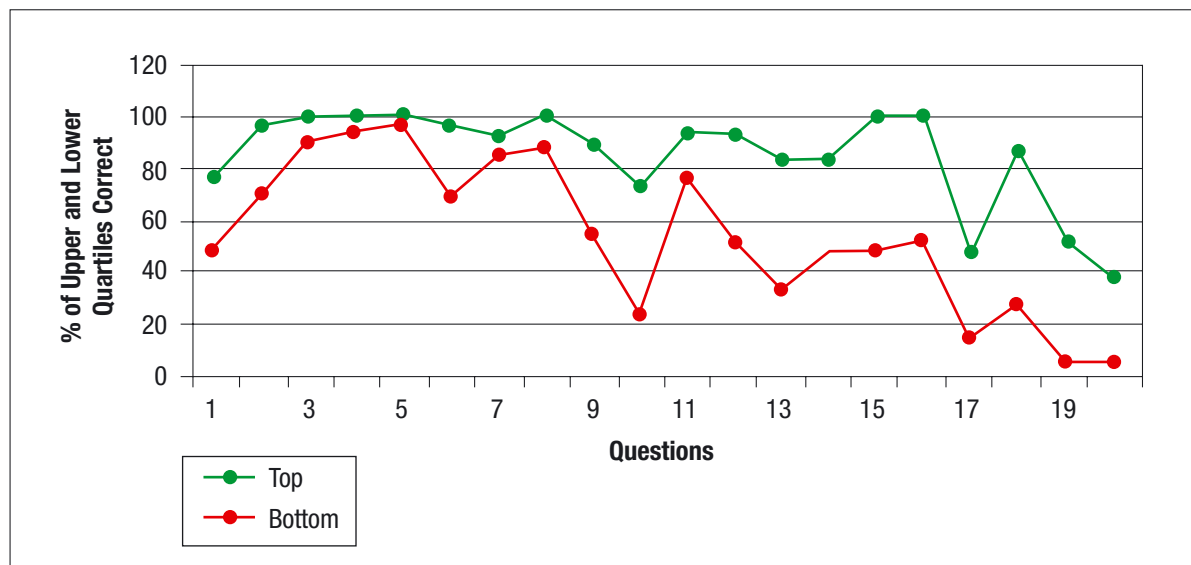
**Figure 2.2: Percentage of the whole class (n=119) answering correctly (Benvenuti & Cohen, 2008)**



MCQs are often criticised for not distinguishing between top and bottom students as success can be achieved via simple rote learning, requiring little more than a surface approach to learning (Scouller, 1998; Woodford & Bancroft, 2005). While this seems a logical argument when the use of MCQs is restricted to testing of knowledge recall, this study sought to examine whether MCQs, specifically designed to test higher level cognitive skills, were capable of distinguishing between top and bottom performers (Research Question 3).

Using Lister's (2005) approach, students were separated into quartiles based on their overall scores for the entire set of questions. Figure 2.3 shows that for knowledge-recall MCQs there is little difference in the performance of the top 25% and bottom 25% of students. However, for higher level questions that require comprehension, application and analysis (questions 10 through 20), the variation in performance is much greater.

Figure 2.3: Comparison of top and bottom performing students (Benvenuti & Cohen, 2008)



Question 10, for example, a comprehension question, enabled strong students to demonstrate their understanding, with 73% of top performers answering correctly. Only 24% of bottom performers managed to answer this question correctly. The application question (question 18) was answered correctly by 87% of the top performers, while only 27% of the bottom performers were able to answer this question correctly. Furthermore, almost no bottom performing students were able to answer the analysis questions (19 and 20) correctly. These results suggest that ‘higher level’ MCQs may be adequate in distinguishing between top and bottom students. A detailed statistical breakdown of the above analysis is given in Benvenuti & Cohen (2008).

## Representative examples of MCQs used in the study

As seen from the above results, several of the MCQs used in the study demanded fairly deep cognitive engagement, and resulted in many students not being able to simply look up the answer from the text. These questions required a deeper understanding of the concepts and neither the availability of the textbook nor unlimited time undermined this. This section provides a brief discussion on some of the questions used in the study (for a detailed explanation of ‘discrimination index’, see Chapter 4 of this volume).

### Question 2

The Internet is a/an

- ☐ a. circuit switching network
- ☐ b. fibre relay network
- ☐ c. \*packet switching network
- ☐ d. optical relay network

Question 2 is an example of a recall question, which required students to simply refer to and recall lecture slides and textbook content. This level of question is most frequently used in MCQ assessment, and these types of questions were generally not good discriminators between top and bottom students (Figure 2.3).

**Question 10**

A branch office of a retail chain needs to upload sales data to the head office. Which one of the following Internet services could be relied upon?

- ☐ a. URL
- ☐ b. \*FTP
- ☐ c. Telnet
- ☐ d. TCP/IP

Question 10 is an example of a comprehension question. Students were required to work with a previously unknown scenario and make a recommendation as to the appropriate Internet service. Top performing students answered correctly by identifying (b) as the correct answer. Many students were unable to understand the difference between the concept of an Internet service and the concept of a network protocol (d) – the answer given by a third of the class. This question is therefore probing for the understanding of a concept often misconstrued by students.

**Question 11**

Search engines are important for Web research because

- ☐ a. They allow a revenue system to be generated from advertising, and this revenue allows the Web to be sustainable
- ☐ b. They add value by having sponsored links delivered with the research results
- ☐ c. \*Users can find the information that they require by searching databases created by the search engines
- ☐ d. Users can search the Web as soon as a new page is created or a website is updated

Question 11 was intended as a comprehension question. However, it was answered relatively well by all students (see Figure 2.3). Post-hoc analysis of this question revealed that the correct answer closely resembled the textbook description of a search engine, which explains the relatively high performance on the question in both top and bottom performers.

**Question 17**

Assertion: Sally wanted to view a hypermedia document (web page) on the World Wide Web, so she needed to type the URL into the address bar of her web browser

BECAUSE

Reason: A web browser is client-side software used to access web pages

- ☐ a. The assertion and the reason are both correct, and the reason is valid
- ☐ b. The assertion and the reason are both correct, but the reason is invalid
- ☐ c. The assertion is correct but the reason is incorrect
- ☐ d. The assertion is incorrect but the reason is correct
- ☐ e. Both the assertion and the reason are incorrect

Question 17 is an assertion-reason question that required students to determine the correctness of each statement, and then to judge whether the reason is an acceptable explanation for the assertion. Students in general battled with this question in comparison to other similarly demanding questions. Upon reflection, the question may have been worded in a manner that made it slightly ambiguous

as to whether the focus of the assertion is the need to use a URL in order to locate a web page, or the need to use a web browser in order to view a web page. This illustrates the importance of wording questions correctly, and the need for close checking and analysis of students' responses following an assessment. Option (a) was selected by 42% of students, possibly due to misinterpretation.

### Question 18

Sally (a travelling salesperson currently in Durban) and Amy (her boss at head office in Johannesburg) want to have a conversation about the sales calls that Sally made that day. Sally has an Internet connection via Wi-Max. Which of the following is the most appropriate choice in the circumstances?

- ☐ a. They should use email because it allows for real-time communication
- ☐ b. They should use instant messaging only if Sally can find another way to connect to the Internet
- ☐ c. They should use VoIP even though it will not allow for real-time communication
- ☐ d. \*They should use video-conferencing because it supports both voice and visual communication
- ☐ e. They should use a podcast because it will allow Amy to ask Sally questions

Question 18 was set as an application question and 40% of students were unable to correctly answer it. Students were required to demonstrate their understanding of various technologies and apply that understanding to the scenario presented. Given all the options and the reasons presented, the most likely course of action for Sally and Amy should be (d).

This required students to either:

- dismiss the other options as incorrect or irrelevant for this particular scenario; or
- dismiss the justification for selecting the option (i.e. they could use instant messaging but that would not require Sally to find another way to connect to the Internet i.e. the student needed to understand that Wi-Max does provide a user connection to the Internet).

The students also needed to draw on their understanding:

- of email as a send-store-retrieve based form of communication that is not real-time;
- that VoIP is real-time; and
- that a podcast allows for playback not conversation.

This question required both knowledge and understanding before the student could attempt to apply that knowledge to make a recommendation.

### Question 19

A student was asked the following question: "Briefly list and explain the various uses of the Web". As an answer, this student wrote the following:

"The Internet is a physical infrastructure that enables a number of services. Without the Internet, the World Wide Web could not exist. Various uses of the Web include: email, FTP, IRC and VoIP. Email accounts for the majority of Internet traffic."

How would you judge this student's answer?

- ☐ a. EXCELLENT (all uses of the Web have been identified with clear and correct explanations)
- ☐ b. GOOD (all uses of the Web have been identified, but the explanations are not as clear as they should be)

- ☐ c. MEDIOCRE (one or two uses of the Web are missing, OR the explanations are not clear OR the explanations are irrelevant)
- ☐ d. \*UNACCEPTABLE (more than two uses of the Web are missing AND the explanations are not clear AND/OR they are irrelevant)

Questions 19 and 20 were both only answered correctly by the strongest students. With question 19, students answering the question incorrectly were evenly distributed across good, mediocre and unacceptable, indicating that students may have been guessing. The correct answer is (d) because:

- 1) all uses of the web are missing (the answer does not identify any uses of the web, such as online shopping, banking, social networking, news, education, health, entertainment, etc.),
- 2) the student in the question has confused uses of the web and Internet services, and has demonstrated a common misconception, and
- 3) although it is a correct statement that the Internet is a physical infrastructure that enables a number of services and that without the Internet the World Wide Web could not exist, none of that is *relevant* to the question posed to the student. It is all background.

Interestingly this type of question highlights student thinking about the manner in which open-ended questions can be answered – ‘dump down everything you can recall about the Internet and hope something fits’. Thirty-two per cent of students felt the answer was ‘good’. This illustrates their lack of understanding of the difference between Internet services and uses of the web, testing a common misconception.

### Question 20

A student was asked the following question: “Briefly list and explain the technologies that make the Internet work”.

As an answer, this student wrote the following:

“The Internet is a physical infrastructure that enables a number of services. Firstly, it is important to recognize that the Internet is a circuit-switching network. This allows for messages to be broken up into packets and dynamically routed to the destination computer. TCP/IP is the communications protocol for the Internet. The protocol defines how messages are broken up into packets, addressed, delivered and reassembled. Routers are computers on the Internet responsible for the forwarding of packets. All computers on the Internet require a unique TCP address. Most Internet services rely on a peer-to-peer computing model. For example, the Web is based on a peer-to-peer model.”

You are reviewing the work in an attempt to help detect the errors and fix the answer.

Which of the options 1–5 represent good corrections to make?

1. The word “circuit” should be replaced by “packet”
2. The word “Routers” should be replaced by “Domain name servers”
3. The term “TCP address” should be replaced with “IP address”
4. The term “peer-to-peer” should be replaced by “server dominated”
5. The word “Internet” should be replaced with “World Wide Web”

- ☐ a. 1 to 5 are all correct
- ☐ b. 1, 2, 3 and 4 are correct but 5 is incorrect
- ☐ c. 1, 2 and 3 are correct but 4 and 5 are incorrect
- ☐ d. 1, 3 and 4 are correct but 2 and 5 are incorrect
- ☐ e. 1 and 2 are correct but 3, 4 and 5 are incorrect
- ☐ f. \*1 and 3 are correct but 2, 4 and 5 are incorrect

Question 20 required students to do four things:

- 1) draw on a number of topics and explanations presented in class around the technologies of the Internet and how the Internet works (a student would need to put a number of puzzle pieces together);
- 2) evaluate and pass judgement on the fictitious student's answer, which requires determining both correct and incorrect statements;
- 3) evaluate each of the possible corrections to determine if they are appropriate; and
- 4) accept that it may not be possible to completely reconcile the answer, i.e. inaccuracies may still remain as some suggested corrections may not improve/correct the student's answer.

The majority of students selected (d), which is a partially correct answer demonstrating that they are unable to fully evaluate the 'correctness' of the recommended correction. The term 'peer-to-peer' should have been replaced with 'client-server' not 'server dominated'. In fact the term 'server dominated' was invented and does not appear anywhere in the students' course notes, or textbook (even a search on Google for the term 'server dominated model' returns no results).

## Reflections on the results of the study

The first research question clearly showed that IS academics are not currently using MCQs to assess students at the higher levels of cognitive demand, specifically at the levels of application and analysis. The vast majority of the questions across the set of exam papers in Research Question 1 were set at the level of knowledge recall. A review of the examination papers revealed that examiners are relying on short-answer or paragraph-style questions to assess higher level skills. The majority of these questions test at the knowledge recall level with few requiring students to demonstrate understanding or application of concepts. These questions generally required students to list, define, identify, discuss and describe, as opposed to explain, distinguish between, predict the outcome of, evaluate, judge or recommend (Benvenuti & Cohen, 2009).

The second part of the study indicates that carefully designed MCQs can be used to tease out common misconceptions, assess higher level cognitive skills, and discriminate between top and bottom students. Moreover, the use of MCQs allows for automated testing and immediate feedback to students on their performance, together with the benefit of reduced marking for assessors.

Several important aspects relating to these types of MCQs should, however, be noted. Firstly, careful consideration must be given to the mark weighting of questions to ensure that students are given enough time for questions that challenge them at higher cognitive levels. As can be seen in the discussion relating to questions 17 to 20 above, each question requires students to do several things, which can include recalling facts; evaluating statements; making judgements; assessing several similar facts, concepts or applications; dealing with common misconceptions; etc. Failure to signal to students the difficulty of the question through a well thought-out mark allocation could mean that students fail to spend the required time working through the whole question and might instead panic, guess at the answer and rush on to other questions.

Secondly, these types of questions make cognitive demands on students that they may not yet have encountered. They require students to work with the bigger picture, sifting through facts, concepts and applications (Biggs, 2003 and Lohman, 1993). Tackling this type of question for the first time in an exam situation would be placing unfair demands on a student. Using MCQs to illustrate and model thinking relating to a particular topic either in tutorials or as part of a lecture helps students to see how facts and concepts have relationships, and how concepts need to be understood and fully unpacked before they can be applied in particular contexts and situations. This approach is being adopted by lecturers across the various topics for the first year IS course.

This brings us to the third issue, that of the need to build a large set of MCQs that allow for use in teaching and assessing, in lectures, tutorials, tests and exams. Lecturers need to plan for the additional time needed to build up a set of questions relating to their topic, and to strategically reuse exam questions in lectures and tutorials in subsequent years. Building an inter-university test bank involving lecturers from the other introductory courses is currently being explored.

Lastly, the need to test the MCQs both before and after use to check for problems with the items is important. Issues relating to ambiguity, language and so on remain as relevant when working with higher order questions as always. Post assessment analysis of both questions 11 and 17 revealed possible problems. Once these problems have been detected, these questions can be improved for future usage in tutorials, tests and exams.

## Conclusion

The reality of large classes requires a strategic response that balances the tensions between reliability, validity and manageability of assessment methods. Carefully designed and well-structured MCQs can be used to address reliability, validity and manageability of assessment. However, the construction of MCQs that assess higher level cognitive skills is not an easy task and requires investment in time in setting questions. The return on this investment, however, is seen in the manageability of marking of assessments and in the detailed feedback that is obtained when analysing the overall performance of a class on specific questions or sections of exams.

The results of this study further show that MCQs can discriminate effectively between top and bottom performers, a long-standing criticism of MCQ testing. The analysis of different types of MCQs shows that this criticism holds true for questions that test student recall, but seems not to hold true for questions requiring more cognitively challenging responses.

## References

- Anderson, WL, 2005. Objectives, Evaluation and the Improvement of Education. *Studies in Educational Evaluation*, 31, 102–13.
- Benvenuti, S & Cohen, JF, 2008. An In-Depth View of MCQ-based Testing in Information Systems. *South African Computer Journal*, 42, 21–28.
- Biggs, J, 2003. *Teaching for Quality Learning at University* (2nd ed.). The Society for Research into Higher Education & Open University Press.
- Brown, G, 2001. Assessment: A Guide for Lecturers. LTSN Generic Centre.
- Brown, G, Bull, J & Pendlebury, M, 1997. *Assessing Student Learning in Higher Education*. London: Routledge.
- Gibbs, G & Jenkins, A, 1992. An Introduction: The Context of Changes in Class Size. In G. Gibbs & A. Jenkins (Eds). *Teaching Large Classes in Higher Education: How to Maintain Quality with Reduced Resources*, 11–22. London: Kogan Page.
- Knight, P, 2001. A Briefing on Key Concepts. LTSN Generic Centre.
- Lister, R, 2005. Methods for Evaluating the Appropriateness and Effectiveness of Summative Assessment via Multiple-choice Examinations for Technology Focused Disciplines. In Proceedings of Evaluations and Assessment Conference (Sydney). Available online: <http://www.iml.uts.edu.au/EAC2005/papers/Lister2005.pdf>.

- Lohman, DF, 1993. Teaching and Testing to Develop Fluid Abilities. *Educational Researcher*, 22(7), 12–23.
- Ramsden, P, 1992. *Learning to Teach in Higher Education*. Routledge, London.
- Rowntree, D, 1987. *Assessing Students: How Shall We Know Them?* London: Kogan Page.
- Scouller, K, 1998. The Influence of Assessment Methods on Students' Learning Approaches: Multiple Choice Question Examination versus Assignment Essay. *Higher Education*, 35(4), 453–72.
- Woodford, K & Bancroft, P, 2005. Multiple Choice Questions Not Considered Harmful. In Proceedings of the 7th Australasian Conference on Computing Education, 42, 109–16.



# A major effect of a minor intervention

## The use of a classification instrument to create awareness of assessment practices amongst lecturers in a first year biology course

*Elisabeth Brenner, Grace Moletsane and Marissa Rollnick*

### Introduction

This chapter discusses an approach to make academics aware of their assessment practices. The strategy has resulted from research that aimed to ascertain and evaluate the assessment practices in a first year biology course at a South African university. More specifically, this investigation was done to explicitly examine whether, in line with the university's Assessment Policy, the course chosen for investigation was using assessment methods tailored to the course content, whether the assessments measured the range of competencies required from first year science students and, finally, whether the quality of the assessments would serve to prepare students for the greater cognitive demands of the second year biosciences curricula. Although the study was carried out on a science course, the approach taken would be equally applicable to any course in which one might wish to inform and create awareness of current assessment practices amongst the lecturing academic staff.

### Background

The introductory biology course was investigated from 2006 until 2008. The rationale for an investigation focusing specifically on assessment practices stems from the notion that assessments may be used to direct students' learning (Gibbs, 2002). Thus, from this perspective, assessment may be considered an integral component of a coherent educational experience (Gravett & Geyser, 2004). This underscores the importance of creating an awareness of the nature of their assessment practices amongst academics. Moreover, in light of the hierarchical nature of the scientific discourse (Bernstein, 2000), it is imperative that assessments are valid and reliable tools to determine whether students have attained the knowledge and cognitive levels needed to access the next stage on the learning path, since knowledge 'gaps' would disadvantage a student's future learning.

The introductory biology course is a full-year course offered at first year level. It is a prerequisite for all of the second year courses offered in the Schools that co-teach the course. The course, which generally has around 400 registered students, is taught on two diagonals to accommodate the large number of students (from different school backgrounds and with different abilities). Teaching is shared between the two participating Schools. The one School, School A, teaches the first semester and this section of the work is examined in June; lecturers from the other School, School B, teach the second semester and this section of the work is examined in November. This means that every student is required to write two 90 minute examination papers in June, one covering material from block one and the second from block two, and two 90 minute examination papers in November, one covering material from block 3 and the other from block 4 (i.e. four papers in total). It was

established that, in the course, all summative assessments are in timed test format (examinations and tests), while the formative assessments, which include practical work and laboratory reports and which are marked by laboratory demonstrators, have a lower status and do not count much towards the year mark as the marking is not considered rigorous enough to be reliable. Furthermore, the questions required for completion of the laboratory reports remain unchanged from year to year, and it seems that they are more often than not copied by students from work submitted by others in previous years. At present there is no laboratory test. In light of this information, it was decided that an analysis of the examination papers set by the two Schools would provide an insight into the standard of the assessment practices in the course.

At the start of the investigation, the focus in both Schools was mainly on *content* taught. Assessment was dealt with independently by the two Schools and questions for formal assessment papers were submitted by individual members of the lecturing staff to the course administrator, who compiled the paper. The result of this was that none of the lecturers were aware of the overall composition of the examination paper (there was no discussion between the lecturers or with the academic coordinator with respect to content coverage, scope or depth of the questions), or of the standard or content of questions other than those that the lecturer had set. Exacerbating the situation in one of the Schools was that block 2 (which was seven weeks long) of the first semester, was generally taught by five different lecturers. This was worrying because, as a core course, it is expected to adequately prepare students for the more rigorous requirements of subsequent courses taken in the second year of study.

Although, for the past four years, approximately 70% of students had managed to pass their first year core biology course, second year results and discussion with the second year lecturers indicated that they had done so without achieving the competencies that would have enabled them to easily manage their second year courses. It was therefore anticipated that an intervention targeting assessment in this course might offer substantial long-term benefits to the students in improving their preparedness for second year, and, since it appears that student performance is determined by what is expected of them in assessments, could potentially change learner attitudes.

## Data collection process

Primarily we examined the written evidence of assessment, namely the examination scripts. Thus, the data collected in this regard were the two examination papers (November 2006 and November 2007) set by School B, and the two examination papers (June 2007 and June 2008) set by School A.

In addition, in order to get a feel for how the course was run, we conducted one interview at the beginning of the process with the course coordinator from School A. Finally, as a result of our findings from analysis of the first examination paper from each School, we conducted an intervention and thus held separate workshops for the staff in each of the Schools.

The papers under study are summarised in table 3.1 below.

**Table 3.1: The papers under study**

Date	Paper number
November 2006 (2 papers)	3
	4
June 2007 (2 papers)	1
	2
November 2007 (2 papers)	3
	4
June 2008 (2 papers)	1
	2

# Analysis

## Analysis tool

An analysis tool for evaluation of the examination papers was developed. This tool, which is shown in Table 3.2, was based on a revised version of Bloom's Taxonomy and a grid modified from Anderson (2005). The modification also took into account the work of Green and Rollnick (2007), which shows that it is not possible to use the whole range of Bloom's Taxonomy or any of its variants in a timed test for science because analysis, synthesis and evaluation in science are processes that need more time than is available in timed tests. In addition, in a scientific work environment these activities would usually be done with additional available resources. Green and Rollnick had therefore suggested that these abilities should be measured in formative tasks rather than in summative assessments. In light of this, Bloom's higher cognitive levels 4 to 6 were 'collectively considered to comprise high order thinking abilities' (Green & Rollnick, 2007:256).

**Table 3.2: Grid used for analysis of questions**

	The Cognitive Process Dimension			
The Knowledge Dimension	Remember (Level 1)	Understand (Level 2)	Apply (Level 3)	Higher abilities (Level 4)
Factual (F)				
Conceptual (C)				
Procedural (P)				

The cognitive level separates between questions that can be answered from recall of knowledge (level 1, remember); questions that require comprehension of concepts (level 2, understand); questions that require application of knowledge to unfamiliar situations (level 3, apply); and questions that require analysis, synthesis and evaluation (level 4, collective higher cognitive abilities).

In addition, questions were also categorised on the knowledge dimension as factual (F), conceptual (C) or procedural (P) after Shavelson (2002). *Factual* questions refer to those where the knowledge can be learned, understood and applied without reference to the underpinning concepts. On the other hand, *conceptual* questions require a deeper knowledge of the underlying concepts. Nevertheless, since they may be answered from recall alone, it is possible to classify a question as C1 (conceptual level 1). It should be noted that *procedural* questions would include plotting of graphs, drawing of diagrams and calculations, and, once again, these operations might be completed by simple recall of previous examples (level 1), require understanding (level 2), require application (level 3) or might require synthesis and/or derivations that would place them at level 4.

## Data analysis

An analysis of the examination papers set by the two Schools was conducted by categorising questions in terms of the above taxonomy. Three evaluators judged each question and classified it according to the grid shown in Table 3.2. Where there were disagreements, these were usually resolved by discussion so that reported results represent the consensus. In addition, the format of each question was recorded and an analysis was done to indicate the composition of the paper with respect to the question type.

## Findings

The analysis of the two November 2006 examination papers is given below. Also included are examples of questions from these papers, justifying their classification, in order to provide an exemplar of how various biological questions may be fitted to our analysis tool.

The first paper (paper 3, written in 2006) examined the work taught in the third block. The paper consisted of 51 questions of different lengths, with mark ratings ranging from 1 to 15 marks. The mark total was 100, to be completed in 1.5 hours. This averages at just over a mark a minute. In the paper, 78% consisted of multiple choice questions, each worth one mark, while 22% consisted of short-answer questions, with the occasional extended answer where a paragraph or more was required. The average number of marks per question was thus calculated at slightly less than 2, indicating that the multiple choice section impacted strongly on the final outcome of the paper. A number of multiple choice questions required a much deeper understanding of the concepts before the correct option could be selected. Only two questions (4%) of the 51 questions were classified at level 4.

Analysis and classification of questions on the cognitive and knowledge dimensions are shown in Table 3.3.

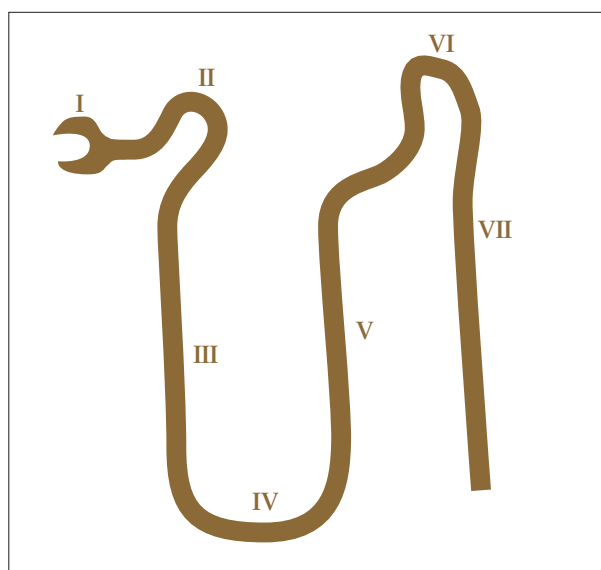
**Table 3.3: Analysis of ILS Biology 111/126 paper 3 November 2006\***

The Knowledge Dimension	The Cognitive Process Dimension				Total
	Remember (Level 1) marks	Understand (Level 2) marks	Apply (Level 3) marks	Higher abilities (Level 4) marks	
Factual (F)	3	2	2	0	7
Conceptual (C)	18	51	4	4	77
Procedural (P)	0	1	15	0	16
Total	21	54	21	4	100

\* Total number of marks = 100

An example of a question from Section B that we classified at level F1 is shown below.

Refer to the figure shown below, a diagram of a renal tubule, to answer the following questions



15. In which region would filtration occur?

- a) I
- b) III
- c) IV
- d) V
- e) VII

F1 (factual – recall) suggests that a student is able to remember which region on the diagram filters urine without necessarily understanding why this is so and without even being aware of what the diagram represents.

On the other hand, question 16 (shown below), which refers to the same diagram, was classified as F2 (factual – understanding), suggesting that a student can answer this question without understanding the process of filtration in the kidney tubule. Unlike question 15, which requires little more than recall of information, this question requires a deeper understanding of the process that leads to urine becoming more concentrated.

16. In which region would urine become more concentrated?

- a) I
- b) III
- c) IV
- d) V
- e) VII

Table 3.4 shows the breakdown of the examination by number of questions, rather than number of marks. This breakdown privileges multiple choice questions, of which there are 40.

We classified three questions at level F1, and two questions at level F2, as shown in Table 3.4 below.

**Table 3.4: Analysis of ILS Biology 111/126 paper 3 November 2006\***

The Knowledge Dimension	The Cognitive Process Dimension			
	Remember (Level 1)	Understand (Level 2)	Apply (Level 3)	Higher abilities (Level 4)
Factual (F)	3	2	-	-
Conceptual (C)	17	21	4	2
Procedural (P)	-	1	1	-

\* Total number of questions = 51

We classified 17 questions out of the paper as conceptual level 1 (C1). An example of such a question is shown below:

**Question 17**

Which part directly leads to the area where membranes become permeable to urea?

- a) III
- b) IV
- c) V
- d) VI
- e) VII

A student who has conceptual knowledge of the different sections of the urinary system is able to recall the different regions.

C2 (conceptual – understanding) refers to questions (21 in total) that require a higher level of cognitive understanding. For example, question 28 (Section B) was afforded this classification:

**Question 28**

Breathing is regulated by

- a) energy intake
- b) the diaphragm and intercostals muscles
- c) somatic centres in the brain
- d) the circulatory system
- e) automatic centres in the brain.

An example of a conceptual question at level 3 (C3, conceptual – apply) requires an even higher level of cognition, in that one should also be able to apply the concept and possibly relate it to other areas. Four questions out of the paper were classified at conceptual level 3. An example is question 2(g).

**Question 2(g)**

How would a drug that interferes with cross-bridge formation affect muscle contraction?

To answer this question, a student must have knowledge of drug interactions and their function, as well as the type of muscular reactions that result from certain drug interactions.

A conceptual question at level 4 (C4, conceptual – higher abilities) requires application of higher cognitive abilities, such as analysis, synthesis and evaluation, as well as relational abilities. Question 2(h), Section A fell into this category.

**Question 2(h)**

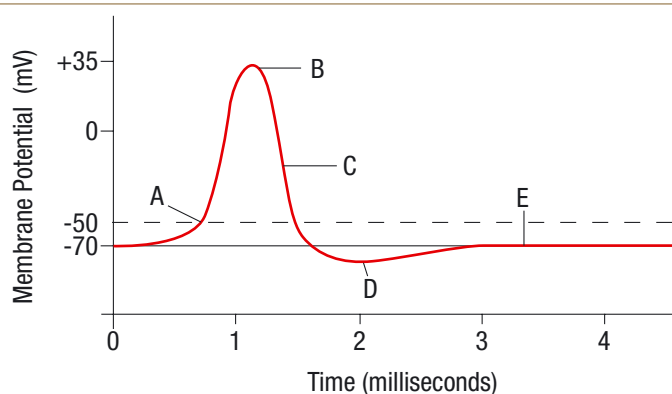
Briefly explain how the functions of the sensory receptors on the tongue and in the nasal cavity determine the flavour of a fruit sweet?

The question is pitched at a higher cognitive level. A student must understand the functions of the sensory receptors on the tongue and also possess the ability to describe the relationship between the nasal and mouth cavities in detecting the flavour of various foods. In this regard, the answer requires synthesis of an explanation from various information areas.

Procedural questions on the knowledge dimension require a degree of understanding that could lead to the application of the knowledge to different situations, such as subject-specific skills, for example in biological drawings, subject-specific techniques and methods, such as plotting and interpretation of graphs or the use of a scientific method. Two questions out of the paper were classified as procedural, one at level 2 and one at level 3. These questions are shown below, with the rationale for their classification.

**Question 8**

For the following questions, refer to the graph of an action potential in the figure below and use the letters to indicate your answer



The threshold potential is reached

- a) A
- b) B
- c) C
- d) D
- e) E

This question requires a student to read the area of maximum potential from the graph. This could be done with limited understanding of how the electrical activity develops in a muscle or nerve cell during a neural impulse. The question was classified at a higher level than simple recall as some understanding of how graphs work is required to answer the question. For this reason it was classified as level 2 (P2). As explained previously, the process of reading and interpreting graphs could be described as procedural as it becomes a routine activity which would be identical in any given situation.

In contrast, question 2 (Section C), which is shown below, was classified as procedural at level 3. This question requires the student to draw an oxygen dissociation curve and explain the change in its shape as a function of the pH of the environment, which requires application of theoretical knowledge. From a content perspective, to answer this question the student must draw the graph showing the variables involved viz oxygen-haemoglobin levels and pH and explain the change in the dissociation curves if the pH of the blood changes or when oxy-haemoglobin dissociates in an acidic environment.

### Question 2

Draw and fully label a graph that shows the Bohr Effect. Explain why the dissociation curve is this shape.

Table 3.5: Analysis of ILS Biology 111/126 paper 4 November 2006\*

The Knowledge Dimension	The Cognitive Process Dimension				Total
	Remember (Level 1) marks	Understand (Level 2) marks	Apply (Level 3) marks	Higher abilities (Level 4) marks	
Factual	10	0	0	0	10
Conceptual	14	18	32	5	69
Procedural	4	7	10	—	21
Total	28	25	42	5	100

\* Total number of marks = 100

The analysis of the questions in paper 4 (2006) shows that, unlike paper 3, where 74% of the marks could be obtained from questions set at levels 1 and 2, over 40% of the marks were from questions set at levels 3 and 4. In paper 4, just over 50% of the marks could be obtained from questions set at levels 1 and 2.

Table 3.6: Analysis of ILS Biology 111/126 paper 4 November 2006\*

The Knowledge Dimension	The Cognitive Process Dimension			
	Remember (Level 1)	Understand (Level 2)	Apply (Level 3)	Higher abilities (Level 4)
Factual	3	—	—	—
Conceptual	7	5	4	1
Procedural	2	2	—	—

\* Total number of questions = 24

Three questions out of the paper were classified as factual. An example of a question classified at level F1 is shown below.

**Question 2b 1**

What would you expect to happen in a year with an abnormally long winter and short summer?

In this question, a student can refer to common knowledge to remember the effect a prolonged winter season may have on the biosphere, without understanding the complex interactions within an ecosystem.

However, 17 questions in this paper were classified as **conceptual**. These required understanding as well as the ability to identify the interrelationships among the basic elements. Seven questions were classified at level C1 (conceptual – remember). Five questions out of the paper were classified at level C2 (conceptual – understand). An example of a question at level C2 is shown below. The first part of the question is at level C1, asking for recall, but in the second part, the student must understand the definition in order to produce an answer. This requires a student to have an understanding of the classification and the categories contained within an ecosystem and also to have the ability to represent them diagrammatically.

**Question 3a**

Define thermal stratification (use a diagram to illustrate your answer). What are the implications of thermal stratification on nutrient cycling in a lake ecosystem?

Four questions were classified at level C3 (conceptual – apply). These questions required students to apply knowledge they have acquired in the course in unfamiliar situations. The question below is an example of a question that we classified at C3.

**Question 4a 1**

Discuss what would happen to the tree–grass ratio (how would the structure of the savanna change) if fires occurred every six months instead of annually?

One question out of the 17 (Section A, question 4a 2) was classified at conceptual level C4 (conceptual – higher abilities). The question required synthesis of an argument working with the effects of various factors on the structure of the savanna. This question is shown below.

**Question 4a 2**

Discuss what would happen to the tree–grass ratio (how would the structure of the savanna change) if the wildebeest population grew rapidly and grazing increased?

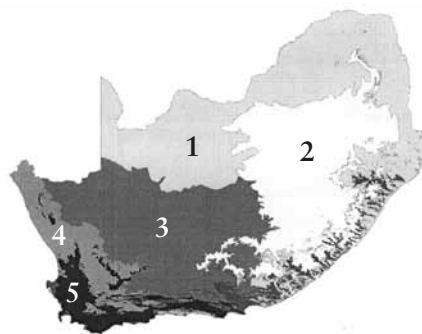
Four questions were found to be in the **procedural knowledge dimension**. The questions required cognitive development at levels P1 (procedural – remember) and P2 (procedural – understand). Examples of these questions are as follows:

## Procedural knowledge dimension at level 1

### Question 2a

Study the map below and answer the questions that follow:

According to Low & Rebelo (1997), there are 7 biomes in South Africa. Give the names of the biomes that do not have numbers on the map above.



## Procedural knowledge dimension at level 2

### Question 2c.

Apart from biome 5, which other biome is threatened by afforestation? Give the name and number of the biome.

To respond to this question the student must know the meaning of afforestation, identify it from the map provided, and also know and understand how the biomes can be affected by this phenomenon, justifying the classification at level P2.

### Summary of the findings of the two November 2006 papers set by School A

Several issues arise as a result of this analysis. Firstly, consideration of the knowledge dimension confirmed that in both papers, the questions are largely conceptual as opposed to procedural and factual. Secondly, referring to the cognitive demands of the papers, it is interesting to note that the predominant level of examination in paper 3 (Tables 3.2 and 3.3) is at the comprehension level, while in paper 4 (Tables 3.4 and 3.5) it is at the application level.

This may have several causes. It is possible that the staff examining different sections may have different styles. Another possible factor may be the nature of the content in the sections covered in each paper, which may not lend itself to deeper questioning. The analysis did, however, suggest that there was a lack of shared vision between the various examiners. In light of this, it was decided to address the issue primarily by making the examiners of School A aware of the discrepancies and to enable them to be more strategic in their design of examination papers in the future. More detail on this intervention is given later in this paper.

### Summary of findings of the two June 2007 papers set by School B

The analysis of the June 2007 examination papers is presented in Table 3.7 which also shows the total marks awarded to the different areas on the knowledge dimension as well as the total number of marks at each cognitive level.

**Table 3.7: Results of analysis of June 2007 paper 1**

	The Cognitive Process Dimension				
The Knowledge Dimension	Remember (Level 1) marks	Understand (Level 2) marks	Apply (Level 3) marks	Higher abilities (Level 4) marks	Total
Factual	13	0	0	0	13
Conceptual	11	60	16	0	87
Procedural	0	0	0	0	0
Total	24	60	16	0	100

In the first paper, 87% of the marks could be obtained from questions set at levels 1 (11%) and 2 (60%), and 16% at level 3. Most of these questions were conceptual (87%). 13% of the marks could be obtained from factual recall. There were no procedural questions in this paper.

Table 3.8 shows the results of the analysis of June 2007 examination paper 2. Once again, the majority of the marks (84%) came from questions in the conceptual knowledge dimension, 13% of the marks were from questions classified as factual (factual – remember) and only 3% of the marks came from procedural questions. On the cognitive dimension there was better balance in this paper in that 31% of the marks came from questions pitched at level 1, 32% from questions pitched at level 2 and 37% from questions pitched at level 3. There were, however, no questions that could be classified at level 4.

**Table 3.8: Analysis of June 2007 examination paper 2**

	The Cognitive Process Dimension				
The Knowledge Dimension	Remember (Level 1) marks	Understand (Level 2) marks	Apply (Level 3) marks	Higher abilities (Level 4) marks	Total
Factual	12	1	0	0	13
Conceptual	19	28	37	0	84
Procedural	0	3	0	0	3
Total	31	32	37	0	100

## The intervention

Intervention workshops were given separately to the lecturers in each School. The purpose of the intervention was threefold:

- to inform the lecturers of the results of the analysis of the question papers and the categories used in the analysis and of their importance in terms of preparation of students for further study in the field;
- to enable lecturers to use the taxonomy to classify questions for themselves; and
- to provide a forum for discussion of formal summative assessments so that each examination paper could, in future, be constructed as an entity that would subscribe to the principles of the course and steer students into what would be expected of them as well as into achieving the outcomes.

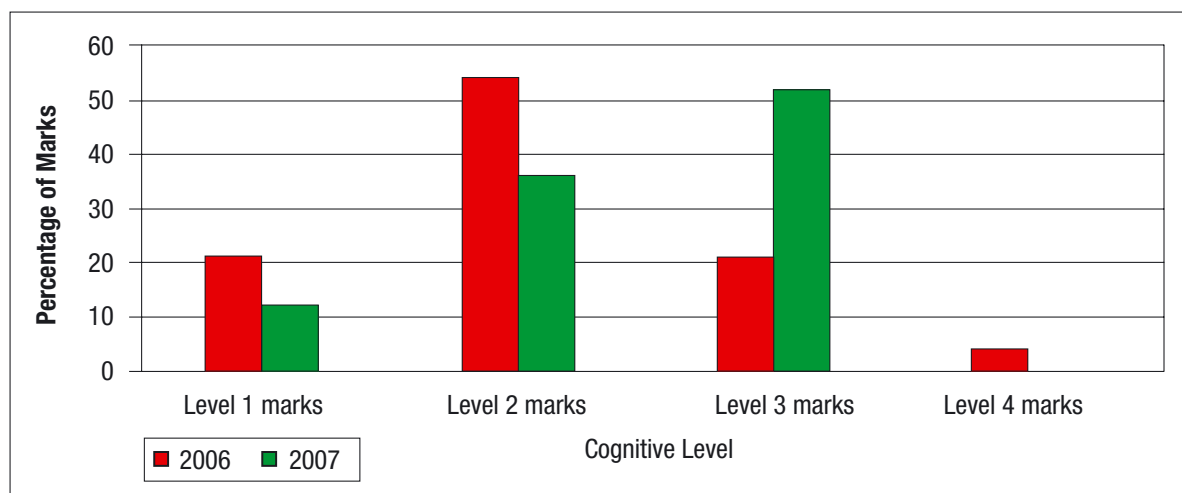
As outlined above, the aim of the intervention workshops was to create awareness of the assessment practices in each School, focusing on constructive alignment between teaching, assessment and desired outcomes. Thus, in the first instance, our assessment tool was introduced and explained to the participants. The results of our analysis of the November 2006 examination papers were also shown. Also discussed was the concept of constructive alignment between teaching and assessment. Lecturers were asked to classify and reach consensus on the classification of some unseen questions in these content areas so that, in future, they could use the analysis tool when setting papers.

All the workshops generated lively discussion and were enthusiastically received. Anecdotal evidence suggested that the knowledge gained by the lecturers on the theoretical aspects of assessment had influenced their assessment practice. It is noteworthy that lecturers from School B met to discuss the compilation of the examination papers. It was reported that this resulted in questions being changed to achieve a greater balance on the different knowledge dimensions and cognitive levels in the papers. To confirm these reports, we analysed the examination papers set by School B in June 2008 and by School A in November 2007, in order to ascertain whether our intervention had had a demonstrable effect on the composition of the examination papers. The results of these analyses compare the question composition of the November 2006 and 2007 papers and the June 2007 and 2008 papers.

## Results of the intervention

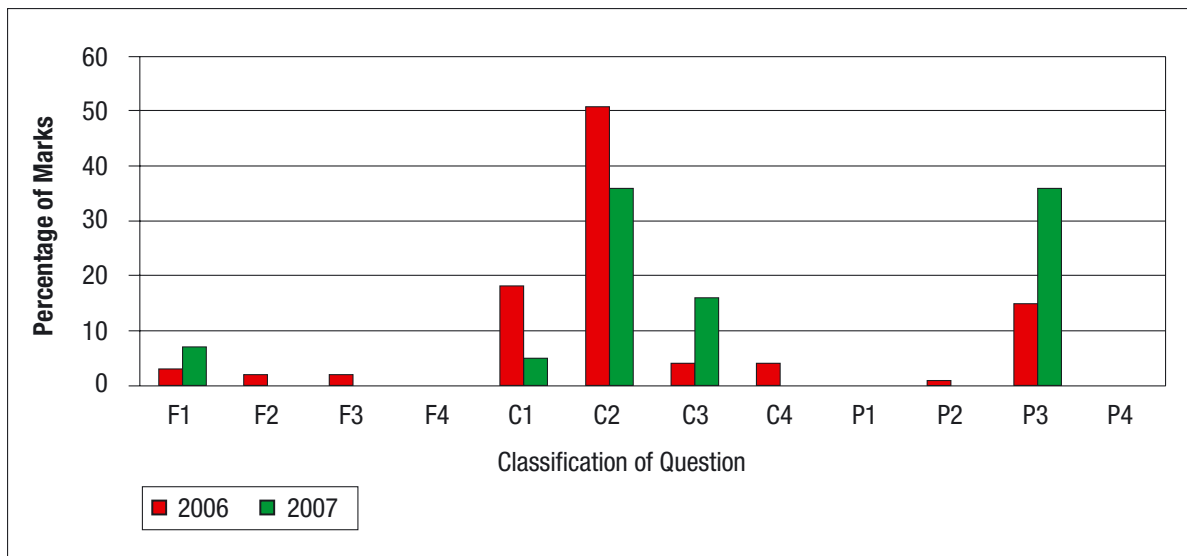
It is evident from Figure 3.1, that there had been a shift to asking questions at a higher cognitive level after the intervention. In 2006, most of the marks (54%) could be obtained from questions asked at cognitive level 2 – understanding. However, in 2007, 52% of the marks could be obtained from questions pitched at the application level – i.e. level 3.

**Figure 3.1: Comparison of marks obtained from questions at various cognitive levels in paper 3 pre- and post-intervention (November 2006 and November 2007)**



Also interesting is that there was an increase in the number of marks that could be obtained from questions in the procedural knowledge dimension in 2007. In the 2007 examination paper, 36% of the marks could be obtained from questions classified as P3, compared to only 15% in 2006. This is evident in Figure 3.2.

Figure 3.2: Comparison of the percentage of marks obtained from questions in different knowledge domains and at different cognitive levels pre- and post-intervention in examination paper 3



As can be seen from figures 3.3 and 3.4, the composition of examination paper 4 did not change vastly from 2006 to 2007. It might be noted, however, that this paper had already been pitched at a higher cognitive level than the other three papers in this course. Moreover, although in the procedural questions content had been somewhat low and remained so, the content of the material that this paper examined did not lend itself to these types of questions.

Figure 3.3: Comparison of marks obtained from questions at various cognitive levels in paper 4 pre- and post-intervention (November 2006 and November 2007)

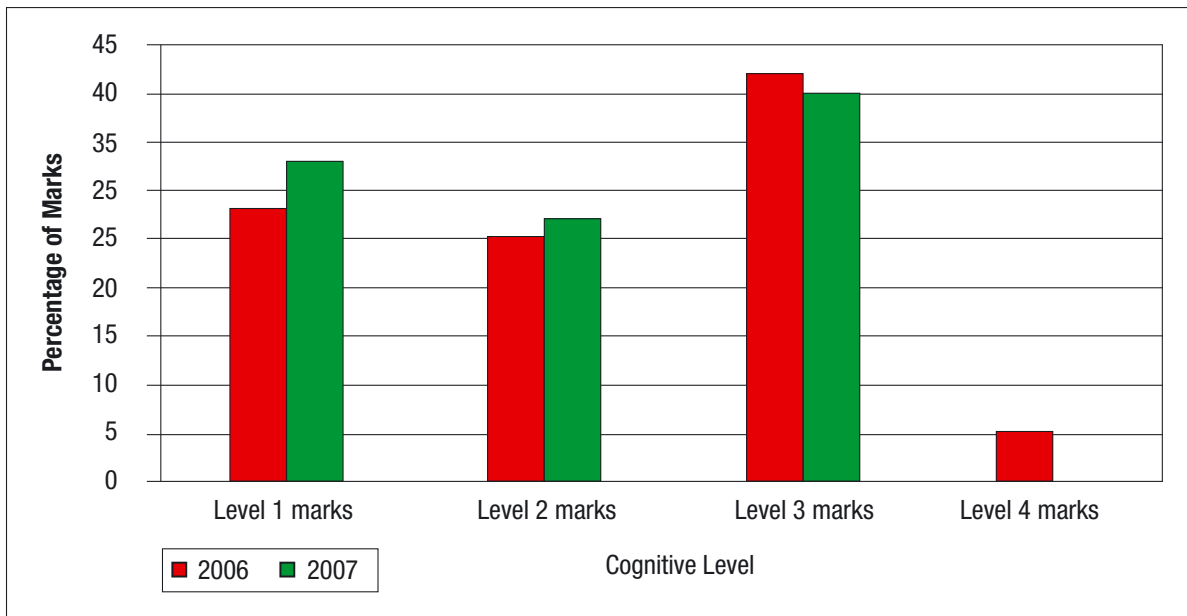
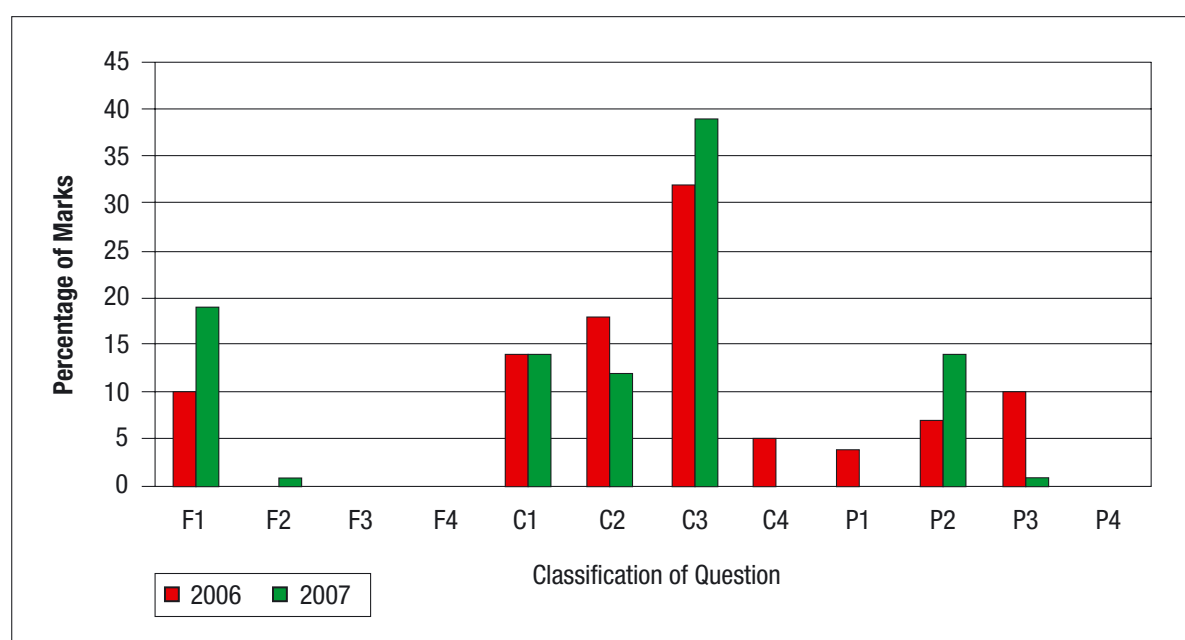


Figure 3.4: Comparison of the percentage of marks obtained from questions in different knowledge domains and at different cognitive levels pre- and post-intervention in examination paper 4



Perhaps the greatest effect of the intervention occurred in paper 1. Table 3.9 shows a comparison of the number of marks that could be obtained from different knowledge areas pre- and post-intervention. While the June 2007 examination paper did not have any procedural questions, 16% of the marks could be obtained from procedural questions in the June 2008 paper. This was at the expense of conceptual questions, as the marks that could be obtained from these dropped from 87% to 69%. Factual questions remained virtually the same (15% in June 2008, compared to 13% in June 2007).

Table 3.9: A comparison of the number of marks that can be obtained from different knowledge areas in the June 2007 and 2008 paper 1

	2007 (marks)	2008 (marks)
Factual	13	15
Conceptual	87	69
Procedural	0	16
Total marks	100	100

Figure 3.5: Comparison of the percentage of marks obtained from questions in different knowledge domains and at different cognitive levels pre- and post-intervention in examination paper 1, which was set by School B

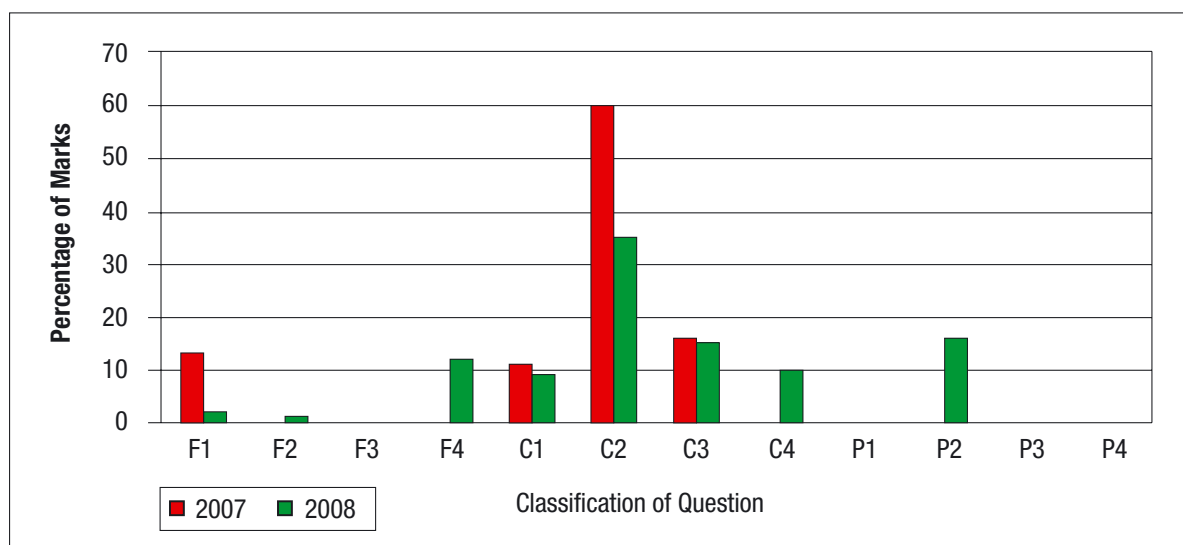


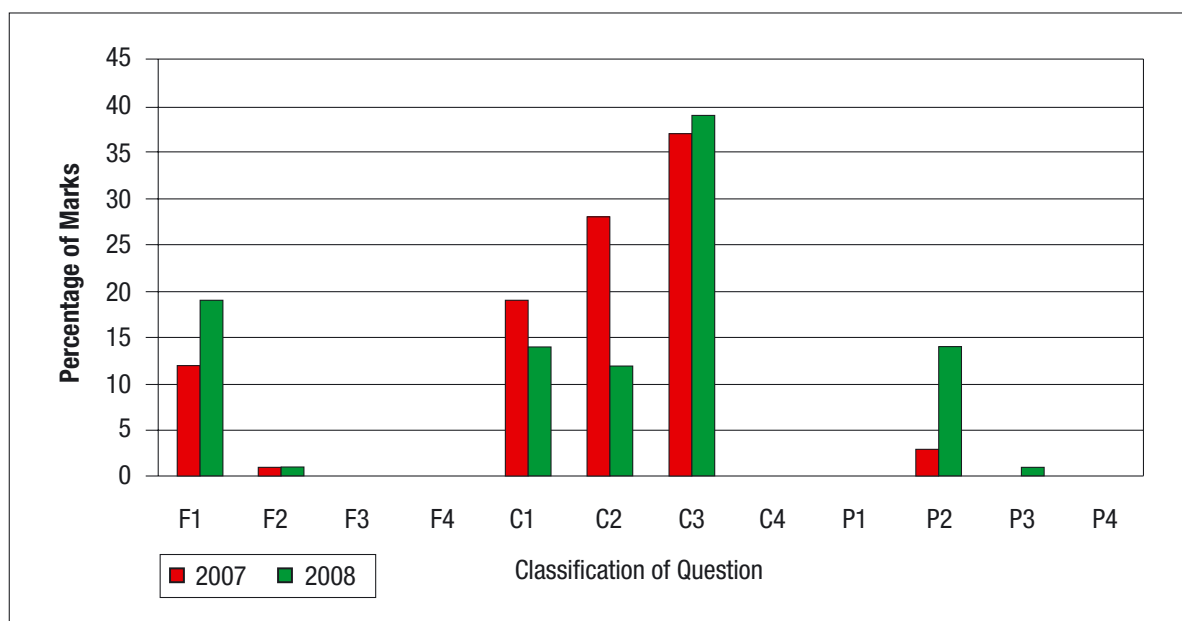
Figure 3.5 shows the comparison of the papers on the different cognitive levels as well. From this it is apparent that the cognitive levels of several questions have increased. In the June 2008 paper, almost 10% of the marks could be obtained from questions classified as conceptual level 4. The June 2007 paper did not contain any questions at this cognitive level. The cognitive level of the factual questions asked had increased from only F1 questions in June 2007 to most at F4 in June 2008. The effect of the intervention on examination paper 1 was thus most gratifying. On a metacognitive level, the inclusion of procedural questions signalled that being a professional in the biological sciences requires one to be able to carry out procedures, understand them conceptually and apply them in different contexts.

While the effect of the intervention was not as pronounced in paper 2 as it had been in paper 1, it is evident from Table 3.10 that the lecturers had made an effort to also include more procedural questions after the intervention since the percentage of marks that could be obtained from procedural questions had increased from 3% in June 2007 to 15% in June 2008. This was at the expense of the lower levels of conceptual questions. As shown in figure 3.6, 38% of the conceptual questions are at level 3.

Table 3.10: Total marks on the different knowledge dimensions in paper 2

	2007 (marks)	2008 (marks)
Factual	13	20
Conceptual	84	65
Procedural	3	15
Total marks	100	100

Figure 3.6: Comparison of the percentage of marks obtained from questions in different knowledge domains and at different cognitive levels pre- and post-intervention in examination paper 2



To sum up, it is interesting that the procedural content of papers 1, 2 and 3 was increased in the second set of papers that we analysed. This was thought to be a direct consequence of the intervention where staff first became aware of the existence of the three categories of questions. As would be expected, conceptual questions were extensively used in examinations and the analysis initially showed a predominance of level 2 questions, also found by Green and Rollnick (2007) in Chemistry in first year. After the workshops, there was also an increase in the cognitive level of paper 1, 2 and 3 questions, which one can again attribute to increased consciousness of the existence of different cognitive levels through the workshop intervention.

Since many academics teaching at universities are not trained in education and hence lack some of the essential tools for design of assessment tasks, it is felt that even a minor intervention such as a short workshop held during lunchtime can be valuable and produce important changes in practice. In general, our experience has shown that interventions such as workshops that are short and to the point are well received and well attended and do have the desired effect on practice.

## References

- Anderson, LW, 2005. Objectives, Evaluation and the Improvement of Education. *Studies in Educational Evaluation*, 31, 102–13.
- Bernstein, B, 2000. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. Oxford: Rowman & Littlefield Publishers.
- Chan, CC, Tsui, MS, Chan, MY & Hong, JH, 2002. Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Students' Learning Outcomes: *An Empirical Study. Assessment and Evaluation in Higher Education*, 27(6), 511–27.
- Gravett, S, & Geyser, H, 2004. *Teaching and Learning in Higher Education*. Pretoria: Van Schaik.
- Gibbs, G, 2002. Evaluation of the Impact of Formative Assessment on Student Learning Behaviour. *Learning Communities and Assessment Cultures: Connecting Research with Practice*. European Association for Research into Learning and Instruction. Newcastle: Northumbria University.
- Green, G & Rollnick, M, 2007. Let's Stop Complaining about Low Cognitive Levels in Testing in Time Limited Examinations: The Case for Fair Testing Practices. *South African Journal of Higher Education*, 21(2), 255–66.
- Shalem, Y & Reed, Y, 2007. *Putting Assessment to the Test*. The Star Series: Education in South Africa Today. Wits School of Education, University of the Witwatersrand, 15–18.
- Shavelson, R, Li, M, Ruiz-Primo, MA & Ayala, CC, 2002. Evaluating New Approaches to Assessing Learning. Keynote Address. Joint Northumbria/EARLI Assessment Conference, University of Northumbria, Newcastle.

# Item analysis as a useful diagnostic tool in the design and analysis of multiple choice questions

*Marissa Rollnick, Elisabeth Brenner and Grace Moletsane*

## Introduction

In the context of assessing large classes at the first year level, a common rational choice made by lecturing staff is to do at least a portion of the assessment through the use of multiple choice questions (MCQs). This choice may lead to good or bad assessment practices, depending on the knowledge the staff possesses about the principles of good design of multiple choice questions and the lessons they learn from administering them.

Detractors of MCQs often claim that the potential for guessing in these items negates their usefulness in assessing learning. Those in favour of MCQs maintain that, with five distractors in a question, the probability of guessing is only 20% and this can be factored into the analysis. Nevertheless, the speedy marking of MCQs is a deciding factor in lecturers' choice of this form of assessment and they are widely used for the assessment of large classes at the undergraduate level.

Good MCQs require time and skill to design and to achieve quality standards requires pre-testing to make sure the questions are free from design faults and discriminate effectively between high performing and low performing students. Most science academics do not possess basic knowledge of item design and an understanding of what makes test items challenging. Those academics who are conscious of their lack of knowledge in this regard opt to use test banks produced by testing agencies or publishers, while academics who are unaware of the skills required, design items and hope for the best.

The study described in this chapter took place across two Schools of biology at a large urban university. Both Schools co-teach a large first year course. Most of the lecturing staff in these schools have little experience in test design but use self-designed MCQ items in their examinations.

The purpose of this chapter is to use some statistics about MCQs collected from the course and show how the statistics can be used to:

- Find links between cognitive level and performance.
- Identify design problems in MCQs.
- Identify misconceptions in the students from their responses.
- Find information about bad/misleading teaching.

## Background

The data collected for this chapter formed part of a larger study which looked at the assessment practices in a large first year biology course. As outlined in chapter 3 of this volume, by Brenner et al., the courses are required to admit large numbers of students from different backgrounds and of disparate abilities because it is a prerequisite for all second year biological sciences courses in the two Schools. The first half of the year is taught by lecturers from the one school (School A) and the second half by lecturers from the other (School B). Each School has a course coordinator to allocate curriculum content to various lecturers from their School. The enrolment in the course varies from 350 to 500.

At the onset of the research there were at least five lecturers from School A teaching during the first half of the year and five lecturers from School B in the second half of the year. This means that students were exposed to 10 different lecturers in the course of the year and examination papers are compiled by the course administrator. There was no discussion between the lecturers or with the academic coordinator with respect to content coverage, scope or depth of the questions.

Assessment in the course (i.e. from both Schools) is in the form of one summative test per block and two summative examination papers at the end of each semester. Questions were set individually by each lecturer and given to the administrator, a non-academic, who compiled the test and exam papers. It was also established that there was little difference between the levels of questions in the test and examinations. On account of this, assessment papers used for analysis in this research project were the summative examination papers. The two papers written during the June examination sessions were set by lecturers in School A, and those written during the November examination sessions were set by lecturers in School B.

## The data used in this study

Table 4.1 shows the location of the MCQs analysed in this chapter, together with their weight in each examination. All papers carried a maximum of 100 marks.

**Table 4.1: Location of MCQs analysed**

Paper	Year	No of MCQs	Number of marks	% of total marks
November paper 3	2006	40	40	20
November paper 4		0	0	
June paper 1*	2007	6	6	12.5
June paper 2		19	19	
November paper 3	2007	21	21	28.5
November paper 4		18	36	
June paper 1	2008	7	7	11.5
June paper 2		16	16	

\*Test statistics not available for this paper

Hence, 121 out of a possible 127 MCQs were analysed, worth 139 out of a possible 145 marks. A few points on this table are worthy of mention.

Firstly, the number of marks allocated to MCQs is not consistent across papers, nor are the possible marks awarded to each question. Although the questions in November paper 4 of 2007 were awarded twice as many marks as any of the others, an analysis of the questions did not suggest that they were any more demanding, either in terms of time or level. The decision to award two marks to each of these questions seems to have been ad hoc. It is also worth noting that School A relied far less heavily on MCQs in the development of the examination paper than School B.

## Analysis conducted

As indicated in chapter 3, use was made of evaluators' classification of the cognitive level of the items. For convenience the grid used is reproduced in Table 4.2.

**Table 4.2: Grid used for analysis of questions**

	The Cognitive Process Dimension			
The Knowledge Dimension	Remember (Level 1)	Understand (Level 2)	Apply (Level 3)	Higher abilities (Level 4)
Factual	F1	F2	F3	F4
Conceptual	C1	C2	C3	C4
Procedural	P1	P2	P3	P4

In addition, we conducted an analysis of the students' responses to the items as follows.

The questions were analysed in terms of the percentage choices per option, the facility value (FV) of the question (the fraction of students who chose the correct answer) and the discrimination index (DI), which is the extent to which a question discriminates between the top and bottom performers in the test. DI is calculated by first dividing respondents into thirds based on the overall score in the MCQ portion of the examination. Then the average score of the analysed item is calculated for the groups of top and bottom performers, and the average scored subtracted. It is calculated as follows.

$DI = (X_{top} - X_{bottom})$  where  $X_{top}$  is the fraction of students in the top third who chose the correct answer and  $X_{bottom}$  is the fraction of students in the bottom third who chose the correct answer. DI can take values between +1 and -1. If the index goes below 0.0, it means that more of the weaker learners got the item right than the stronger learners. Such items should be discarded as worthless. In fact, they reduce the accuracy of the overall score for the examination (<http://cms.nelc.edu.eg/help.php?module=quiz&file=itemanalysis.html>).

There is a relationship between FV and DI that is a consequence of the way they are calculated. Clearly, if all students either get an item correct (FV = 1) or all get it wrong (FV = 0), then the item does not discriminate and, as a consequence, DI will be 0. Also, maximum values of DI can be obtained when the FV is between 0.33 and 0.66 (Melton, 2002). Generally, DI values below 0.3 are considered poor. It should also be borne in mind that if the MCQ has five options, the chances of guessing correctly are 20%, so any FV below 0.2 indicates that the value may have been achieved by chance.

## Analysis of questions

### Overall data

#### 1. Relationship between FV and DI:

Figures 4.1 and 4.2 show the relationship between the facility value (FV) and discrimination index (DI) on two of the papers.

Figure 4.1: Relationship between FV and DI on June 2007 paper 2

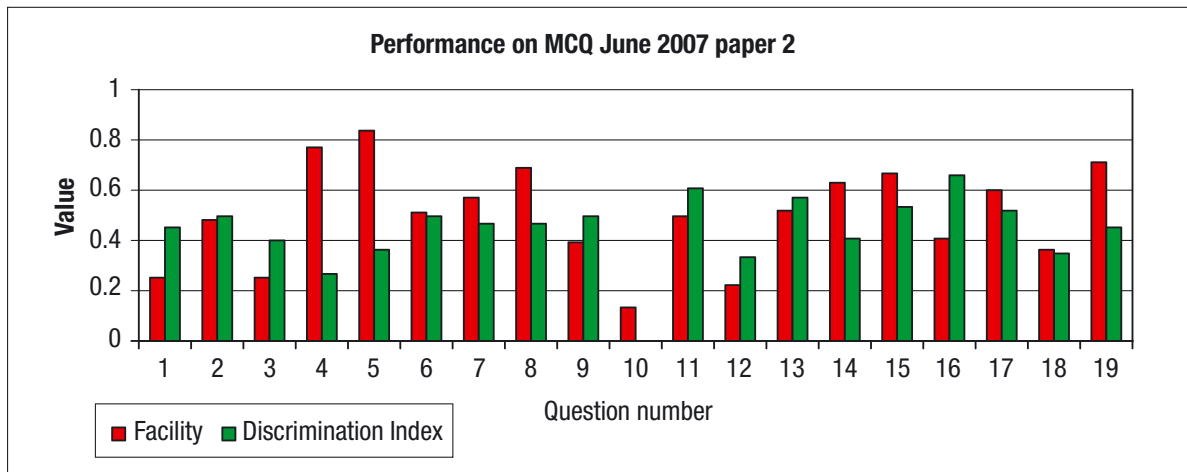
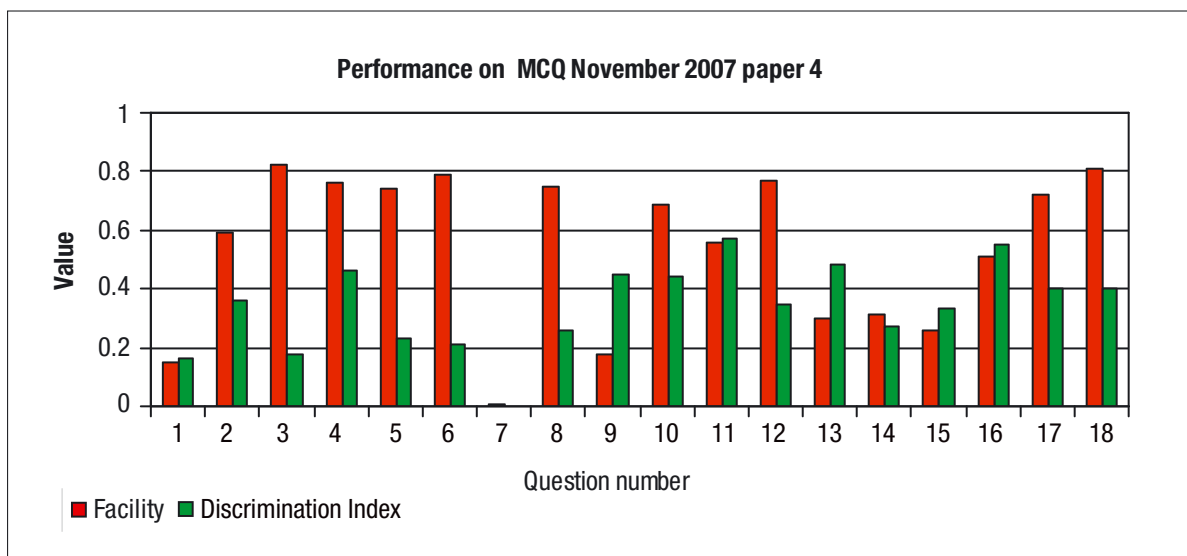


Figure 4.2: Relationship between FV and DI on November 2007 paper 4



As can be seen, the data confirms the point made above, that very high and very low facilities (such as questions 4 and 10 in Figure 4.1 for example) have poor discrimination indices associated with them. This is to be expected, since if all students answer correctly or all answer incorrectly, the question will not discriminate between high and low performers. On the other hand, questions with moderate facility values, like questions 3 and 13 in Figure 4.2, do show good discrimination. This is not always the case. For example, see questions 14 and 15 in Figure 4.1. In such cases analysis of the questions can yield design weaknesses.

Paper 3 in November 2007 stood out in that the entire set of questions discriminated very poorly – this will be explored further below.

## 2. Relationship between cognitive levels as determined by evaluators and students' performance on examinations:

In general, there was not a clear relationship between the cognitive level of the questions as determined by evaluators and the performance of the students on those questions – average facilities for questions at each level were either stable or showed a slight increase with the level of the questions. However, an interesting trend was observed in that the questions classified at a higher cognitive level tended to have a stable or higher average DI. In other words, they discriminated better between the students. This is to be expected, as an able student would be expected to get a higher level question correct while a less able student is more likely to get it wrong. Figures 4.3 to 4.7 illustrate this trend.

So, for example, the average DI for a level C3 question is sometimes higher than the average DI for a level C2 question, as can be seen in figures 4.6 and 4.7, while in Figure 4.5 the average DI for a level C2 question is higher than that of a level C1 question. This trend is difficult to generalise as there were generally not many level 3 questions in the papers – in fact, in Figure 4.5 there are none.

Figure 4.3: Average FV and DI for November 2006 paper 3 in relation to cognitive level

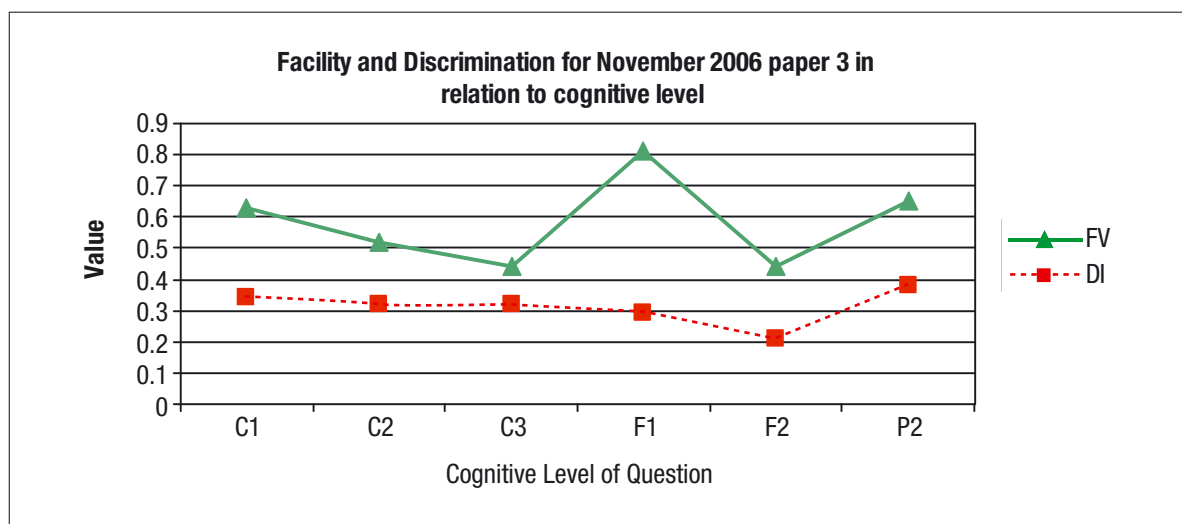


Figure 4.4: Average FV and DI for November 2007 paper 3 in relation to cognitive level

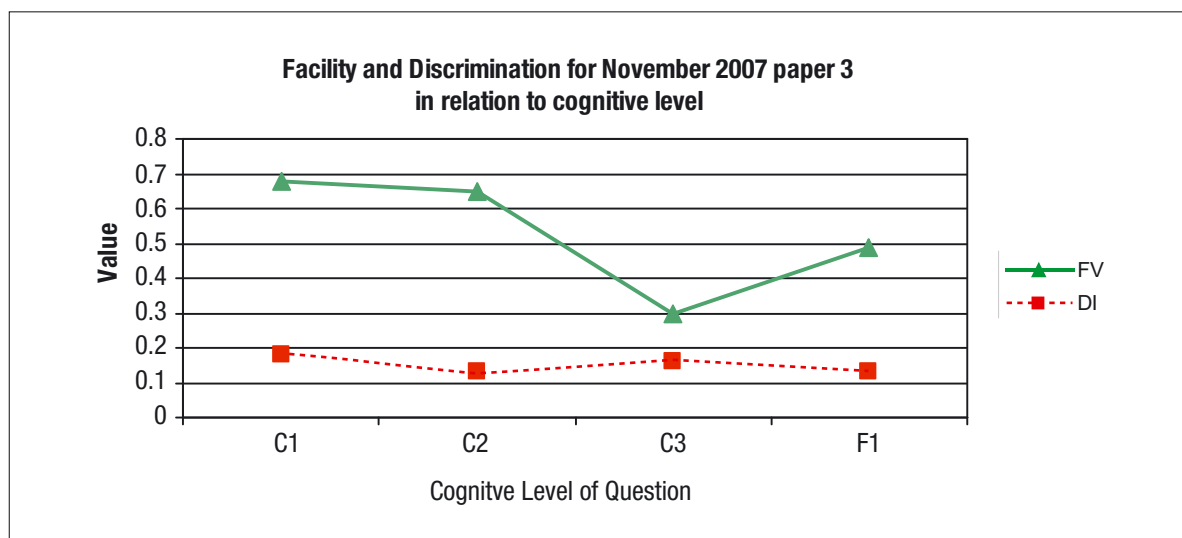


Figure 4.5: Average FV and DI for November 2007 paper 4 in relation to cognitive level

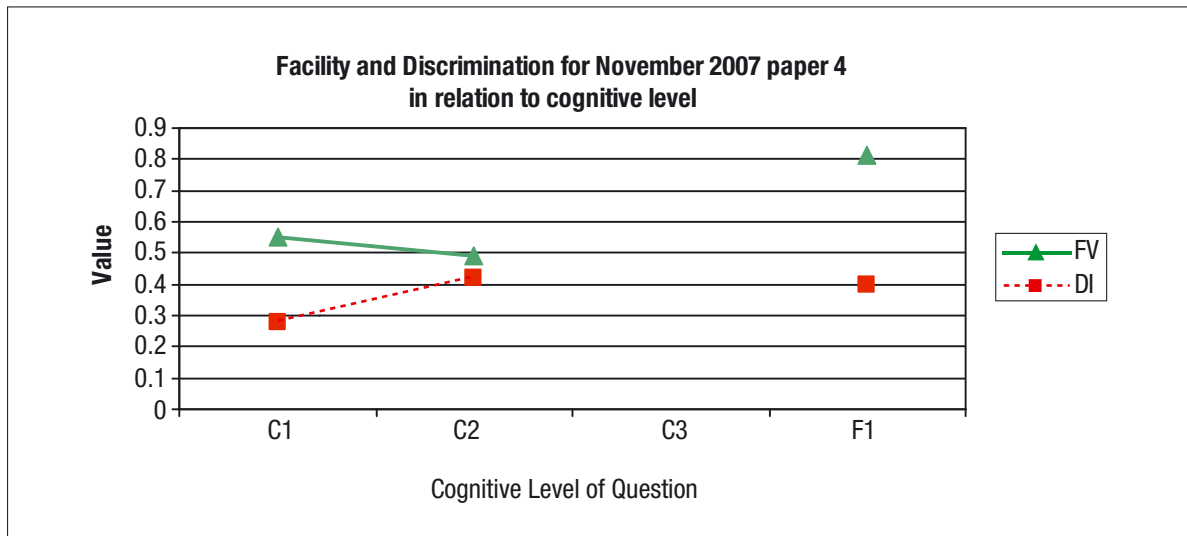


Figure 4.6: Average FV and DI for June 2007 paper 2 in relation to cognitive level

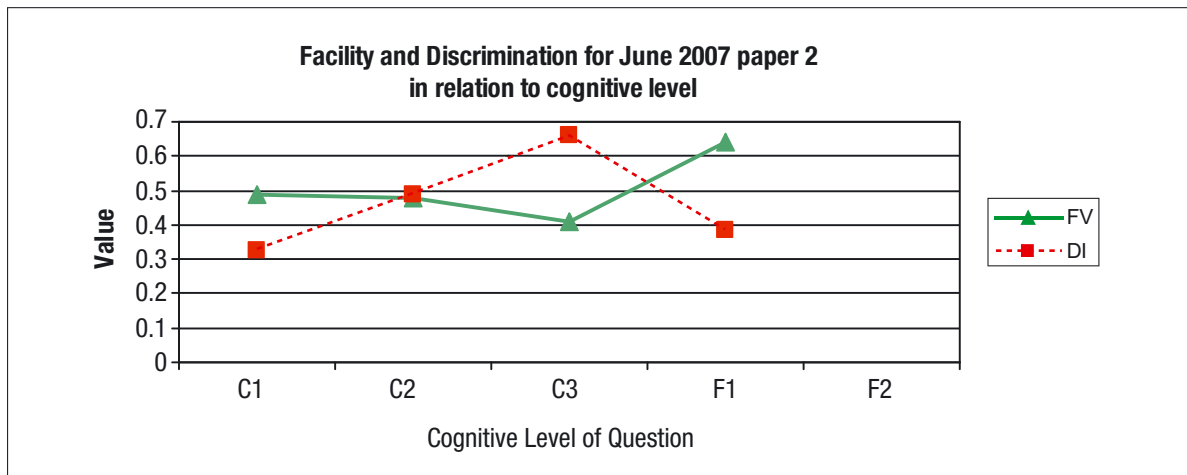


Figure 4.7: Average FV and DI for June 2008 paper 1 and 2 in relation to cognitive level

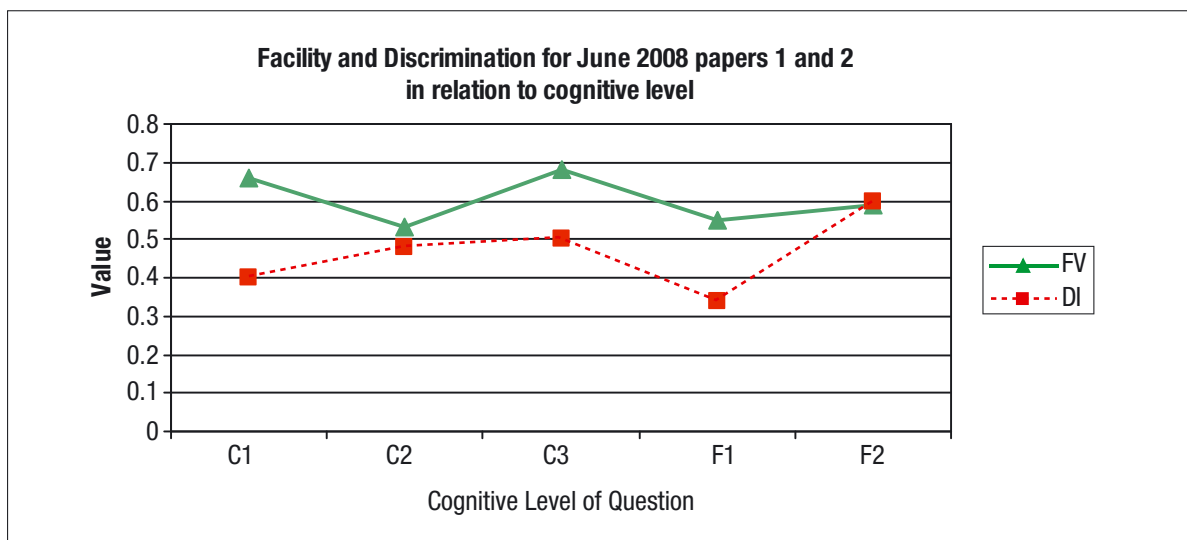


Figure 4.4 stands out in that all 21 questions showed very poor discrimination. The highest DI in the paper was 0.2. Effectively this means that the MCQs in the paper were not discriminating between strong and weak students, identifying a general weakness in all the questions. The facility values, though in some cases high, were within a reasonable range. Twelve of the questions were classified by evaluators as level 1 (F1 and C1) and only one question was classified at C3. The correct options were reasonably well distributed between options A to E, meaning that there were no identifiable patterns that encouraged guessing. All the questions were verbal, some requiring careful reading of a large amount of text, but none requiring either numerical or graphical analysis. Some of the questions selected for analysis below suffered from design faults, but no single reason could be found for the low facility over all the questions. Eleven of the 21 questions had FVs above 70%, which could explain the low DI in those cases.

We now look at questions flagged by their FV and DI values and the kind of issues identified.

## Analysis of individual questions

### Badly and well designed questions

As mentioned above, questions with high and low FV will generally have low DI because of the way that the DI is calculated. This does not mean they are poor questions, but by their nature they do not discriminate because most candidates get the answers correct in the case of high FV and incorrect in the case of low FV. However, in the case of moderate to low FV (0.25 to 0.75), badly designed questions may be identified due to their low DI, as well a poor distribution of distractors, e.g. in question 10 of June 2007 paper 2, shown below.

10. Which of the following is evidence for the occurrence of evolution?
- a) \*Similarities in DNA and protein sequences in various species
  - b) Dinosaur footprints
  - c) Embryonic gill slits in vertebrates
  - d) Drug resistance in viruses
  - e) All of the above

The statistics for this question are shown in Table 4.3.

**Table 4.3: Statistics for June 2007 paper 2 Question 10**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a*	b	c	d	e			
10	13.5	4.9	6.8	12.2	62.4	0.54	0.13	0

\* Shows the correct answer

This is an example of a poorly structured question. Only 13.5% of students answered correctly and the question does not discriminate between high and low performers, as shown by the low DI. The evaluators had difficulty classifying it according to cognitive level. It was classified as C1, then C2 and, before that, F1. The response pattern also shows the problems associated with providing the option 'all of the above'. Logically, if E is the answer, then all the options should be marked correct. Very often test designers resort to the use of 'all of the above' and 'none of the above' when they cannot find credible alternatives to the correct answer. Interestingly, the students in this case did consider E as a viable option, as most of the candidates found it attractive. Nevertheless, the low facility does explain to some extent the poor discrimination.

Question 12 on the November 2007 paper 4 had similar characteristics but shows the effect of combining the options, ‘none of the above’ and ‘all of the above’ into a single question.

12. Which factors most affect nutrient availability in tundra soils?
- a. amount of faeces decomposing
  - b. \*lemming population size
  - c. amount of vegetation cover
  - d. all of the above
  - e. none of the above

The statistics for this question are shown in Table 4.4.

**Table 4.4: Statistics for November 2007 paper 4 Question 12**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a	b*	c	d	e			
12	9.8	1.4	8.6	78.2	2	0.3	0.01	0

\* Shows the correct answer

Again ‘all of the above’ is an attractive option. The low response to option E shows the weakness of including the option ‘none of the above’, which attracted only 2% of the responses. The designer clearly could find no further options after C and caused students to miss option B as the correct answer. The evaluators rated this question at level C1.

Question 16 below, taken from the June 2008 paper 2, is an example of how the statistics highlight a well designed question.

16. Why was the cloning of ‘Dolly’ considered a major scientific breakthrough?
- a) It was the first time a surrogate mother was used successfully
  - b) It was evidence that DNA methylation regulates gene expression
  - c) It showed that cells can be arrested in the cell cycle
  - d) It proved that the pattern of gene expression is controlled at transcription
  - e) \*It showed that differentiated adult cells of mammals can dedifferentiate

The statistics for this question are shown below in Table 4.5.

**Table 4.5: Statistics for June 2008 paper 2 Question 16**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a	b	c	d	e			
16	15.1	17.3	6.8	8.6	48.1 *	4.32	0.48	0.72

\* Shows the correct answer

The question has a moderate FV of 0.48 but discriminated very well at 0.72. The options are all similarly structured, allowing for the candidates to concentrate on the content, rather than the structure of the options. All the options attracted more than 5% of the candidates and the high DI shows that high-scoring candidates answered correctly while low-scoring candidates did not. The evaluators rated this question a C1.

## Misconceptions

Another way in which MCQ analysis can assist in teaching is through the identification of misconceptions. Smith, DiSessa and Roschelle (1993) argue that working with misconceptions can be productive in the learning of accepted scientific ideas. In this way, questions can be diagnostic, and serve as a starting point for learning, as well as evaluative. Questions revealing misconceptions can usually be identified when a particular incorrect option attracts a high level of responses. Question 16 from June 2007 paper 2 shows an instance of this:

16. Within a few weeks of treatment with the drug 3TC, a patient's HIV population consists entirely of 3TC-resistant viruses. How can this result best be explained?
- a) HIV has the ability to change its surface proteins and resist vaccines
  - b) The patient must have become reinfected with 3TC-resistant viruses
  - c) HIV began making drug-resistant versions of reverse transcriptase in response to the drug
  - d) \*A few drug-resistant viruses were present at the start of treatment, and they increased as a part of the viral population soon after treatment started
  - e) The drug caused the HIV RNA to change

The statistics for this question are shown in Table 4.6.

**Table 4.6: Statistics for June 2007 paper 2 Question 16**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a	b	c	d*	e			
16	9.5	1.6	44.3	41.1	3.2	0.81	0.41	0.66

\* Shows the correct answer

One structural weakness of the question is that options C and D are longer than the others and, according to Table 4.6, are equally attractive to the candidates. The question discriminates well, suggesting that more of the lower scoring candidates are selecting option C and higher scoring candidates selecting option D. The evaluators considered the question challenging and classified it at level C3. The question provides useful information to the course instructors, as C may be a misconception which could be taken into account in teaching. However, the statistics also suggest that the question may be improved by working on the other distractors.

Question 22 from the November 2006 paper 3 also highlights misconceptions. The question asks about the circulatory system in the fish as follows:

22. Which one of the following acts as the circulatory system's pump in the fish?
- a) Vascular system
  - b) \*Heart
  - c) Circulating fluid
  - d) Gills
  - e) Vein

The statistics for this question are shown in Table 4.7.

**Table 4.7: Statistics for November 2006 paper 3 Question 22**

Question Number	Option chosen							
	a	b*	c	d	e	Blank	Facility	Discrimination Index
22	7	59.4	5.2	27.6	0	0.3	0.28	0.24

In this question, which discriminates poorly, D emerges as a major distractor, possibly due to the role the gills play in the respiratory system in the fish. Once identified, this misconception can be taken into account in teaching. The statistics do also encourage the course instructors to search for better distractors for the other options which are not attractive. E is such a poor distractor that not a single student chose it. The evaluators classified this question as level C1.

Question 4 from paper 1 in June 2008 is a good example of a conceptual question that identifies misconceptions from both instructors and the students.

4. A catalyst .....
- a) allows an endergonic reaction to proceed more quickly
  - b) provides the activation energy so that a reaction can proceed more quickly
  - c) \*allows an exergonic reaction to occur
  - d) is destroyed once the reaction is complete
  - e) lowers the overall amount of energy needed for the reaction to proceed

The statistics for this question are shown in Table 4.8.

**Table 4.8: Statistics for June 2008 paper 1 Question 4**

Question Number	Option chosen							
	a	b	c*	d	e	Blank	Facility	Discrimination Index
4	10.9	38.8	40.3	6.5	2.4	1.18	0.4	0.5

\* Shows the correct answer

As the question stands, there is in fact no completely correct answer. C may be correct but is incomplete as it stands as exergonic reactions may occur spontaneously anyway provided they have low enough activation energy. B is close to correct but the action of a catalyst is actually to lower the activation energy rather than provide it. A also identifies a misconception that students may have and nearly 11% of students were attracted by this option. E is a well-known misconception at school level but does not seem to prevail at this level. B and C are by far the most attractive options and can open up useful issues for discussion on the conceptual issues around catalysts. The evaluators classified this question as level C2.

### Information about teaching

Closely linked to the issue of misconceptions are questions where other teaching issues may be identified, such as poor student response to questions where answers are spread across a wide range of options, suggesting that the students may not have understood the material at all, as opposed to having a particular misconception, for example question 12 taken from the June 2007 paper 2.

12. Given a population that contains genetic variation, what is the correct sequence of the following events, under the influence of natural selection?
1. Differential reproduction occurs
  2. Mutations occur
  3. A new selective pressure arises
- a) \*2, 3, 1
  - b) 3, 1, 2
  - c) 2, 1, 3
  - d) 3, 2, 1
  - e) 1, 3, 2

The statistics for this question are shown in Table 4.9.

**Table 4.9: Statistics for June 2007 paper 2 Question 12**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a*	b	c	d	e			
12	21.9	19.7	23.8	24.6	9.7	0.27	0.22	0.33

This question has a low FV and relatively low DI. Responses are evenly distributed across the options. This may point to general difficulty for students in this section. It is possible that the different format of this question may have disoriented the students. It is the only question in this format on the paper and could easily be converted to a standard MCQ format. This question was classified as level C1.

Another interesting issue to emerge from the analysis is the difference in statistics on two very similar questions, 34 and 35 from November 2006 paper 3, reproduced below.

34. What function do beta receptors have in regulating the blood flow?
- a) They are stimulated by the sympathetic nervous system, causing vasodilation
  - b) They are stimulated by the sympathetic nervous system, causing vasoconstriction
  - c) They are stimulated by the parasympathetic nervous system, causing vasodilation
  - d) They are stimulated by the parasympathetic nervous system, causing vasoconstriction
  - e) \*They are stimulated by circulating catecholamines, causing vasodilation
35. What function do alpha receptors have in regulating the blood flow?
- a) They are stimulated by the sympathetic nervous system, causing vasodilation
  - b) \*They are stimulated by the sympathetic nervous system, causing vasoconstriction
  - c) They are stimulated by the parasympathetic nervous system, causing vasodilation
  - d) They are stimulated by the parasympathetic nervous system, causing vasoconstriction
  - e) They are stimulated by circulating catecholamines, causing vasodilation

The statistics for this question are shown in Table 4.10.

**Table 4.10: Statistics for November 2006 paper 3, questions 34 and 35**

Question Number	Option chosen					Blank	Facility	Discrimination Index
	a	b	c	d	e			
34	16.8	11.5	30.1	17.5	*23.4	0.7	0.23	0.37
35	23.1	*40.5	9.1	23.1	3.1	1	0.41	0.46

The interest in the response patterns of these two questions is that they differ in only one word in the stem but produce totally different statistics. Both show reasonable DIs but very different FVs and distribution of answers. Another interesting aspect of this question is that it is testing two aspects of a phenomenon simultaneously. In four out of the five responses, the examiners are working with a combination of two alternatives on each. For the fifth response a third alternative is introduced as one of the stimuli. This third alternative contains the correct response in question 34. The evaluators classified both questions as C2, testing conceptual understanding.

Question 35 was an easier question, as demonstrated by the FV, with a higher DI, showing that a higher percentage of better performing students could answer the question correctly. Students easily eliminated option E, realising that catecholamines play no role in the function of alpha receptors, but there is no clear pattern of choice of wrong answers after that to guide teaching, as the popular incorrect responses, A and D, show a combination of involvement of parasympathetic and sympathetic nervous systems with effects of vasodilation and vasoconstriction. There could be an element of guessing responses A to D after students have eliminated E, but the 40% choice of B is above chance.

Question 34 caused more problems as only 23.4% of students chose the correct response and the rest of the responses are distributed among the other options, with nearly a third of the students selecting C. A, C and E make up 70% of the responses, suggesting that these students realise that beta receptors cause vasodilation, but are not certain of the stimulus. In question 35, 63% of the students realise that alpha receptors cause vasoconstriction. It should also be realised that only two of the five responses involve vasoconstriction, so the students who are guessing may complicate this response pattern.

One final issue related to this question is one of technical design. As constructed, it requires a great deal of reading and possible confusion. If the examiners had obeyed one of the rules of MCQ design, the questions would have been far more readable, as follows:

What function do alpha (Q34)/beta (Q35) receptors have in regulating the blood flow? They are stimulated by

- a) the sympathetic nervous system, causing vasodilation
- b) the sympathetic nervous system, causing vasoconstriction
- c) the parasympathetic nervous system, causing vasodilation
- d) the parasympathetic nervous system, causing vasoconstriction
- e) circulating catecholamines, causing vasodilation

## Conclusion

There is more to MCQ design than most academics realise and many elementary design errors were encountered in the analysis of these questions. These include the use of options such as ‘none of the above’ and ‘all of the above’, the repeating of identical phrases in the options, the use of non-equivalent distractors and over use of the negative in the stems of questions, often followed by a negative in one of the options. These poor design features are often more of an obstacle to second language speakers and may affect their performance on MCQ tests.

Since the use of MCQs is widespread in large classes, it is worth putting effort into the design of good questions as much can be gained through the use of these questions at different times, by storing them in item banks. Item statistics are useful in identifying such questions. Tested items are available from other countries, such as the United States, and it may be a good idea to utilise these as they may suffer from fewer design errors. However, they would still need to be tested on South African students.

Since MCQs are scored electronically, minimal effort is required to obtain simple test statistics and, as can be seen from the analysis above, useful information can be gained even from a cursory look at these regarding design problems and student misconceptions, as useful pointers for teaching. The analysis of the last two questions shows that changing only one word in a question can lead to completely different response patterns.

A further point of interest is the lack of correlation between the level of questions as classified by evaluators and the performance of students on these questions, though it appears that the higher level questions do discriminate better. As Brenner et al. show, in chapter 3 of this volume, workshops assisted in making lecturers aware of the existence of the cognitive level of questions and consensus between the researchers and the lecturers on these levels was obtained during the workshops, but the analysis above shows that these levels do not necessarily relate to student performance. This lack of relationship may suggest that these questions had been previously used in class tests or tasks. As explained by Brenner et al. novelty is an important consideration when deciding on cognitive level.

Test statistics are easy to obtain and can reasonably easily be used to extract useful information about the quality of the items, teaching effectiveness and student thinking. With just a simple understanding of the basic statistics, much can be learnt about the success of teaching and the test items themselves can often be used as useful teaching tools in formative assessment during teaching. Equally, items with poor statistics can be eliminated or altered. Relatively little effort thus yields a rich return and potentially an important impact on teaching.

## References

- Learning Content Management System. Available online: <http://cms.nel.edu.eg/help.php?module=quiz&file=itemanalysis.html>. Retrieved 30 January 2010.
- Melton, RF, 2002. *Planning and Developing Open and Distance Learning: A Framework for Quality*. London: Routledge.
- Smith, JP, DiSessa, AA & Roschelle, J, 1993. Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *The Journal of the Learning Sciences*, 3(2), 115–63.



# Immediate response to mediated learning: The use of technology for continuous assessment in higher education

*Elisabeth Brenner and Yael Shalem*

## Introduction

The sciences rely heavily on hierarchical attainment of core knowledge and principles. Informal classroom continuous assessment has traditionally been limited to posing questions to the class, trying to elicit answers from the students, and attempting to gauge from the responses whether sufficient transfer has taken place before moving on. While this approach can give an indication to an experienced teacher as to whether more explanation is needed for the majority of the class, it has been difficult to track individual students' learning and understanding in large classes. Instructors usually have to wait until assignments and tests for formative assessment purposes, and/or summative tests or examinations have been marked before being able to provide feedback, by which time it is very often too late to rectify shortcomings and correct misconceptions. To circumvent this situation and to be able to perform continuous and diagnostic assessments, and thus track individual learning on an ongoing basis, an innovative technology using personal response radio transmitter keypads was introduced into a large second year molecular biosciences class at the University of the Witwatersrand. This course, which is a co-requisite for all the major courses offered in the school that offers it, aims to teach the skills and core knowledge required in the major courses and so it is important that students master the material as it is taught.

In this chapter, we report on the implications of continuous assessment and immediate feedback, which was made possible by this technology, on student performance. We also reflect on the constructs that can be assessed using the keypads and comment on the effect that immediate feedback has on students' acquisition of scientific concepts.

## Content and organisation of knowledge

One of the many challenges facing educators in the tertiary education sector is the expectation that they will facilitate the transformation of novice learners into experts in their field of study. This is not an easy task, especially if one considers that 'Experts and novices are distinguished not just by the content of their knowledge stores, but by their organization' (Beatty & Gerace, 2002:751).

Elaborating this concept, it becomes evident that experts possess a large store of richly connected knowledge within the domain and that they are capable of integrating multiple representations of knowledge constructs, solving problems by applying concepts and can perform qualitative analyses.

Novices, on the other hand, possess limited, disconnected knowledge constructs and, in the sciences, solve problems purely by manipulating equations, generally independently of concepts (Dufresne et al., 2007). In effect then, to become an expert in a particular field requires deep rather than superficial learning, as well as an ability to connect and use key concepts.

This means that, in the transformation process, students have to acquire a broad content base, learn to problem solve by understanding and applying the underlying concepts, and make connections between different content areas. Most importantly, this kind of achievement requires a paradigm shift from the students themselves. Students who were able to achieve high grades in high school, where it was often enough to be a passive recipient of recalled information, are now expected to become active participants in the learning process, to take responsibility for their own learning, and, in terms of the Piagetian developmental model that was applied by Biggs and Collis (1982) to describe the structure of the observed learning outcomes (SOLO taxonomy), to construct knowledge at a relational level. This is the crux of what is required to make the shift from surface to deep learning.

It is evident that many students, in light of their previous experiences in formal education, arrive at university expecting to be able to approach their learning as they always have. A study by Mumba et al. (2002) on the gap between high school and first year chemistry at the University of the Witwatersrand established that while content areas were often the same, high schools valued demonstration of content acquisition, particularly if it was pertinent to what was going to be assessed in the final matriculation examinations, whereas the university stresses the application of content knowledge. Despite this change in emphasis, not all first year students appear to appreciate the cumulative nature of concepts, and resort to the rote learning approach they applied in high school. This means that even if they manage to scrape the minimum passing grade, they enter their second year of study with large gaps in their knowledge with respect to the core knowledge and principles they were expected to have mastered.

Knowledge structures in the sciences have been classified by Bernstein (2000) as having a hierarchical structure. This means that its contents are structured vertically, and each procedure is related to another in a vertical continuum. Theory in hierarchical knowledge structures develops through integration of formal propositions and results in a pyramidal structure of growing levels of complexity. The implication for acquisition is that attainment is achieved only when built up from first principles. However, it must be emphasised that, even though hierarchical knowledge systems lend themselves to systematic progression of acquisition, this does not mean that there is one set of steps per discipline or that there is only one way of structuring the teaching of scientific knowledge (Muller, 2007:82). Muller argues further that in a formal learning environment it is necessary to stipulate a curricular continuum: 'Agreeing that we cannot stipulate a once-and-for-all path, we could still have to concede, retrospectively considered, that there are a specifiable necessary minimum set of steps that must be pedagogically traversed'.

In a subject where conceptual knowledge is cumulative, knowledge gaps severely impact on and restrict students' future learning. From a pedagogical perspective, this means that it is essential that lecturers revisit previous content in order to ensure that essential gaps are minimised and, ideally, eliminated. For purposes of instruction, science courses' curricula must follow a logical sequence, which is determined both by what knowledge the knower has acquired and, in terms of cognitive demands, what the knower is expected to achieve. This means that when selecting learning activities that promote cognitive processes, it is necessary to make forward and backward references, particularly as concepts take a long time to be formed (Dufresne et al., 2007). This supports the idea that the central function of teaching is to pave the way for future knowledge acquisition by making forward and backward references in order to associate new concepts with knowledge that has been, at least, partially acquired. This idea portrays pedagogy as the task of 'making knowledge interwoven and interconnected, rather than linear' (Dufresne et al., 2007:5).

Interconnectivity is indicative of deep learning. A study investigating the interactions between various pedagogic characteristics in teaching science knowledge to children aged between nine and ten substantiates this claim: ‘When going back systematically to scientific knowledge already explored, to relate them with new contents to be learned, the teacher increases the learning time for the former’ (Morais & Pires, 2002:16).

Also, positively influencing scientific knowledge acquisition is a high level of cognitive demand. This was found to be crucial to obtain a high level in science achievement (Morais & Pires, 2002:16).

## Assessment and knowledge

Shavelson et al. (2002:10–11) have defined science achievement as consisting of four types of knowledge. These are:

- Declarative knowledge or factual knowledge, which they explain as ‘knowing that’.
- Procedural knowledge, which is being able to perform a series of steps, or ‘knowing how’.
- Schematic knowledge, which enables one to interpret problems, make predictions, troubleshoot, and is ‘knowing why’.
- Strategic knowledge, which is ‘knowing when, where and how’ to use certain types of knowledge in a situation not encountered previously.

Shavelson et al. have modified the assessment triangle of Pellegrino et al. (2001:4) to develop a conceptual framework for observing and evaluating the results of assessment practice in the form of an ‘*interactive assessment square*’ (Pellegrino et al., 2001:5), which links the knowledge construct with the assessment tasks and design, observation of student performances and interpretation of results with respect to what can be said about the learner.

The above framework for assessment suggests that, as educators and academics are faced with the task of monitoring students’ progress in the process of transformation from novice to expert, they can do this by placing emphasis on the evidential quality of our judgement (Killen, 2003). This idea suggests the use of assessment as a diagnostic tool, on the basis of which lecturers construct a teaching plan to bridge the gap between what they make available to students for acquisition and what students actually take up or grasp. As academics, it is suggested, lecturers are expected to change their approach to assessment from one where assessment is directed at past learning, and thus is applied at the end of a block of learning, to one where assessment is integrated into teaching and is used to inform evaluative judgement during the teaching process (Madaus et al., 1997; Gipps, 1999; 2004; Shepard, 2000; Dann, 2002; Black & William, 1998; Hargreaves, 2005 and Broadfoot, 1999). In essence, this approach argues for the need to shift assessment away from its classical function of quality control and to gear it towards a quality assurance function:

What we need is a shift from *quality control* in learning to *quality assurance*. Traditional approaches to instruction and assessment involve teaching some given material, and then, at the end of teaching, working out who has and hasn’t learned it – akin to a quality control approach in manufacturing. In contrast, assessment *for* learning involves adjusting teaching as needed while the learning is still taking place – a quality assurance approach. Quality assurance also involves a shift of attention from teaching to learning. The emphasis is on what the students are getting out of the process rather than on what teachers are putting into it. (Leahy et al., 2005:19; emphasis in the original)

Shepard (2000) explains the role of assessment in a learning culture in terms of the interlinking tenets of curriculum, constructivist learning theories and assessment strategies. Thus, following a social-constructivist approach to learning, Shepard proposes that assessment practices connect directly to ongoing instruction and be used formatively in support of student learning. Dann (2000:45) suggests that this shift in the role of assessment implies the following:

If learning, which is identified as involving complex structuring and restructuring of information through interactions and experience, is to advance, the interface between assessment and learning must inevitably be dynamic, complex and collaborative.

Lecturers are expected to socialise students into academic criteria, especially if these students are the first generation to have access to tertiary education, and if their ‘social and cultural capital’ has not prepared them for academic practice well enough:

Learning responsiveness of the curriculum entails teaching and assessing students in ways that are accessible to them. This includes making available what is valued about the underlying discipline, how it is assessed and which evaluative criteria are of significance, but also teaching to the rhythms, and the tensions and emotions of learning. (Slonimsky & Shalem 2006:36–7)

Interactive teaching strategies have been found to motivate students to engage with knowledge constructs (Beatty, 2004), so that, as an end point, they can be guided into taking responsibility for their own learning. Active learning is brought about by students constructing knowledge through their own experiences rather than simply reproducing knowledge that was transmitted to them. However, although there are many documented strategies for promoting active learning, the task of assessing the level of the knowledge constructed remains complex (Newmann et al., 1996).

It has also been found that the use of continuous feedback to students (Hirsch & Gabriel, 1995; Taras, 2001; Rust et al., 2003; Higgins et al., 2002 and Knight, 2004) is an important strategy for motivating and enabling students to meet expectations. Research advocates a dialogic feedback between teacher and students, one in which students can recognise their misconceptions (Butler, 1988). Immediate feedback is one form of dialogic feedback, one in which lecturers inform students’ further steps in their acquisition, by identifying the construct they struggle to learn. The question that is often raised is what dialogic type of feedback would mean in the large lecture halls of higher education. Until recently, it has been difficult to find ways to track individual students’ understanding *during* lectures and give immediate and *individual* feedback. In a large lecture theatre, despite a lecturer’s best efforts to engage all students and adjust teaching frameworks and instruction methods in response to observation of student learning, it remains very easy for individual students to disengage, so that lecturers are, in effect, only teaching according to the responses of a small sector of the class. Compounding the situation is that weak students are intimidated and often do not even attempt to answer questions posed to the class. Lecturers therefore generally have to wait until assignments and tests have been marked. In a system whereby individual feedback is delayed, lecturers do not have access to the kind of diagnosis proposed by Shavelson et al. (2002). They are unable to design teaching interventions, which could indicate to students their misconceptions and knowledge gaps *in time* for repair before summative assessment takes place. This is particularly relevant for teaching the sciences, which, as pointed out, rely heavily on a hierarchical attainment of core knowledge and principles.

## Interactive class response systems as a pedagogic tool

Beatty et al. (2006) advocate an ‘interactive classroom response system’, which they believe can enable lecturers at all levels and with any class size, to diagnose and track students’ grasp and understanding during instruction. A recent technological development in the commercial educational sector, the *InterWrite PRS* technology, which uses personal response radio transmitter keypads (‘clickers’), was introduced into a large second-year molecular biosciences class. *InterWrite PRS* technology requires the insertion of various types of assessment questions into PowerPoint presentations on material as it is taught. The assessment types include multiple choice questions, True/False responses to statements, or calculations where answers are keyed-in, thus providing

opportunities to test what Shavelson et al. (2002) classify as declarative, schematic, procedural and strategic knowledge in science. These types of questions are assessed as follows. Procedural knowledge is assessed by getting students to perform a calculation, which they enter on a keypad. Schematic, declarative, procedural and strategic knowledge are assessed by using multiple choice or True/False questions. Individual student responses are transmitted to a radio-receiver connected to the lecturer's computer, which is linked to a projector, and class results are projected onto a screen after each question.

Where appropriate, lecturers set up the questions so that individuals are informed (on the keypads) whether they have keyed in the correct answer as they respond to each question. In the more difficult questions, students are given the opportunity of changing their responses if they answer incorrectly. Moreover, everyone is able to view the responses of the class as a whole. This enables lecturers to respond by adjusting the teaching framework to pace the lecture in line with all students' needs in the class. Lecturers are also encouraged to ask open-ended questions in which they canvass students' opinions in terms of making predictions (schematic knowledge). Asking open-ended questions creates contexts for class discussions, encourages critical thinking and productive engagement with the subject matter, and helps to diagnose conceptual problems in students' understanding. Beatty et al. (2006:32), who argue for a classroom response system of teaching, propose a 'question-driven instruction' model of teaching. They explain that every question used in class should 'serve an explicit pedagogic objective', in terms of the content it wishes to emphasise, the cognitive processes it seeks to evoke, and the metacognition it can bring about.

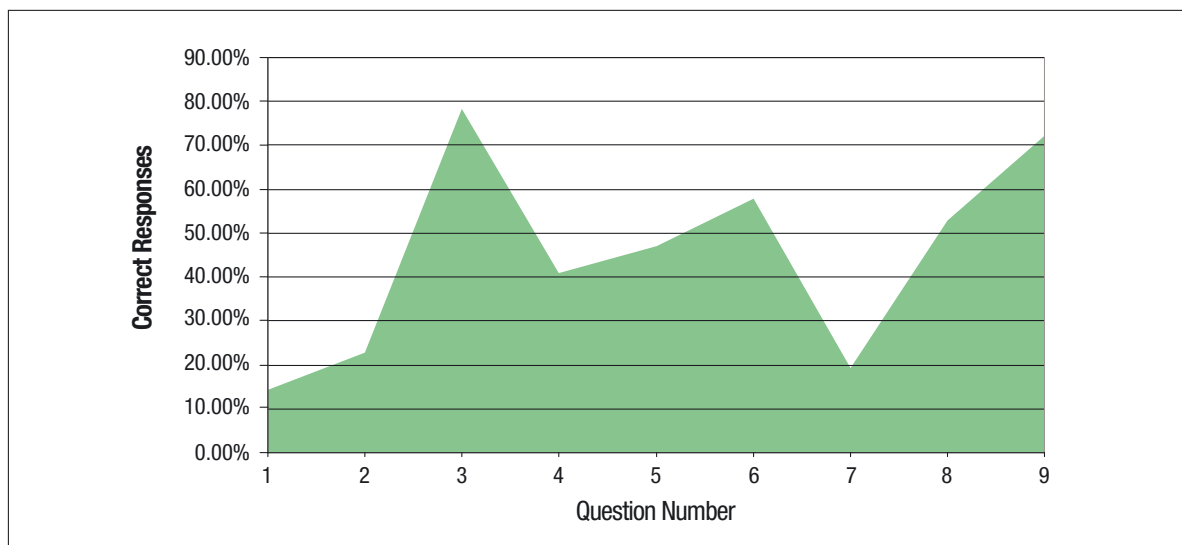
The course, in which the *InterWrite PRS* technology was introduced, is a co-requisite for all the major courses offered in the three teaching programs offered in the school and it is presented to 110 students. The main aim of the course is to introduce students to the skills and core knowledge required in the major courses. It requires students to be able to perform certain mathematical operations and to possess a strong understanding of chemical principles. These foundations are often lacking when students arrive in second year, which makes it imperative for the lecturer to fill in these knowledge gaps.

This paper examines initial findings that emerge from a pilot study and its aim is to investigate the use of this innovative technology, designed to enable ongoing diagnostic (formative) assessment in the classroom. The central claim about assessment that is embedded in this technological innovation is that, with careful question design (Beatty et al., 2006), the assessment device can identify whether concepts have been grasped and whether students are able to perform key operations *as the lecture is taking place*. Feedback is also immediate. The study poses the following research question: *In what ways has this diagnostic form of assessment, used in a large class, informed students' understanding of instruction and supported their subsequent performance on objective tests?*

Student performance during lectures was analysed by investigating their responses to selected questions posed to the class. The study investigated the extent to which subsequent performance in objective summative assessments was supported by immediate feedback given to students during the lecture. Results on a selected question were compared between two groups of students: those who attended the lecture in which immediate feedback was given and those who did not. All students had been exposed to that particular knowledge construct in a practical session.

The content chosen for evaluation reflects a procedural knowledge construct (Shavelson et al., 2002) of how to apply the Henderson–Hasselbalch equation. This is a content area that past students in the course have found particularly difficult to master. It was also felt that the implications of the use of immediate feedback in teaching a procedural knowledge construct could be more easily evaluated. The questions posed during the lecture requiring keypad responses were also analysed in terms of their knowledge constructs. Shown in Figure 5.1 below is a representation of the class responses during the lecture on the application of the Henderson–Hasselbalch equation.

Figure 5.1: Responses to questions asked during the lecture on the Henderson–Hasselbalch equation. The graph indicates the % of correct responses to each question.



The Henderson–Hasselbalch equation relates pH to the  $pK_a$  of weak acids in terms of the ratio of a weak acid to its conjugate base. Questions 1 and 2, which are shown on Figure 5.1, require students to demonstrate that they understand the *concept* of  $pK_a$ . This concept was taught the previous week. Past experience has shown that while students are generally able to give a satisfactory definition along the lines of: ‘The  $pK_a$  is the pH at which a solution of weak acid and its conjugate base exist in equal concentrations’, they are not able to apply the concept or to understand how the Henderson–Hasselbalch equation would verify their definition.

At the start of the lecture, students’ conceptual understanding was assessed, by posing to the class the following two multiple choice questions:

**Question 1:**

What is the net charge on an acetic acid molecule at a  $pH < pK_a$ ?

1. +1
2. +1/2
3. -1/2
4. -1

**Question 2:**

What proportion of acetic acid is ionized at a pH value above the  $pK_a$ ?

1. >50%
2. 50%
3. 0%
4. <50%

Initially, only 15% understood the concept. A variety of responses were obtained from the class, thus affording insight into students' various levels of cognition. This enabled the lecturer to make a judgement on students' understanding of how pH affects the charge on weak acid groups. Although there did not appear to be consensus in understanding this concept, the second question was posed without any lecturer-led discussion to give students time to reflect and self-correct. However, at this point, as is evident in Figure 5.1, it was ascertained that only just over 20% of the class was able to *apply* the concept of  $pK_a$ . This was an important realisation, which necessitated a class discussion of the concept before the Henderson–Hasselbalch equation could be introduced to the class. As the lecture progressed, students were required to insert figures into the equation in order to calculate the  $pK_a$  when given the pH and the weak acid to base ratio (question 3).

$$pH = pK_a + \log \frac{[A^-]}{[HA]}$$

**Question 3:**

What is the  $pK_a$  of a weak acid HA if a solution containing 0.1 M HA and 0.2 M  $A^-$  has a pH of 6.5?

This type of question requires very little in the way of conceptual understanding but requires elementary procedural knowledge. As can be seen in Figure 5.1, almost 80% of the class was able to perform this operation. The 20% of the class that could not do it may lack very rudimentary mathematical skills.

A more complicated level question involving the application of the Henderson–Hasselbalch equation was then presented to the class. The first question asked students to calculate the ratio of weak acid to conjugate base. The second question asked students to calculate the pH from information about the relative concentrations of weak acid and base. These questions require an ability to interconvert an amount of a substance (as number of moles or grams) and concentration (in % or M). Although this is required of first year students, and had been reviewed in an earlier lecture, it has been the lecturer's experience that this is something that many students continue to find difficult to access (some still at third year level). Questions 4 and 5 were therefore inserted at this point in the lecture. These questions elicited correct responses from less than 50% of the class. This indicated a lack of mathematical skills, as well as an inability to perform a fundamental operation in chemistry. For example, question 5 relies on students' ability to relate the formula for density to the one used to calculate the mass of a substance from the number of molecules. Many students in the class were unable to make the connection between the two formulae. This suggests a gap in a strategic knowledge construct (Shavelson et al., 2002), and in terms of the SOLO taxonomy (Biggs & Collis, 1982), students' inability to deal with relational level data. As these are fundamental procedures, time was spent reviewing them.

**Question 4:**

Calculate the amount (g) of sodium acetate to be weighed when preparing 100 ml of a 2 M solution of sodium acetate ( $M_r = 82 \text{ g/mol}$ )?

**Question 5:**

Calculate the amount (ml) of glacial acetic acid you'll need to measure out to make 100 ml of 0.2 M glacial acetic acid ( $M_r = 60.05 \text{ g/mol}$ ; density =  $1.05 \text{ g/ml}$ )

Thereafter, students were asked to calculate the weak acid to base ratio (question 6, using the Henderson–Hasselbalch equation, which once again requires little more than elementary procedural knowledge). They were asked to do this in order to build up their confidence, before they were assessed on their ability to integrate the many types of procedural knowledge that had been assessed and responded to at previous stages in the lecture.

**Question 6:**

If you calculate the proportion of acetic acid molecules that will be dissociated when pH 5.06 (pKa for acetic acid = 4.76) the answer will be:

1. 2/1
2.  $\frac{1}{2}$
3. 1/5

The ratios indicated above were specifically chosen in order to diagnose if students were aware of which figure applied to the acid, and which to the base, as this is a mistake commonly made when students first apply the Henderson–Hasselbalch equation.

The last questions took the application of the Henderson–Hasselbalch equation to its completion:

**Question 7:**

If you had prepared 2 M solutions of acetic acid and sodium acetate, what volume (ml) of sodium acetate would you use to prepare 100 ml of 2 M acetate buffer with a pH of 5.06?

Initial attempts by the class to integrate the many types of procedural knowledge, indicated (question 7) that only 20% of the class was successful. Problems were encountered when students tried to convert the ratio of the weak acid to base into volumes of each of the components. The students' inability suggested a conceptual gap with respect to the mathematical concepts of proportion and fractions. The lecturer's explanations of these concepts increased the success rate to 53% of the class (question 8), and, by the end of the lecture, to over 70% (question 9).

**Question 8:**

If you had prepared 2 M solutions of acetic acid and sodium acetate, what volume (ml) of acetic acid would you use to prepare 100 ml of 2 M acetate buffer with a pH of 5.06?

**Question 9:**

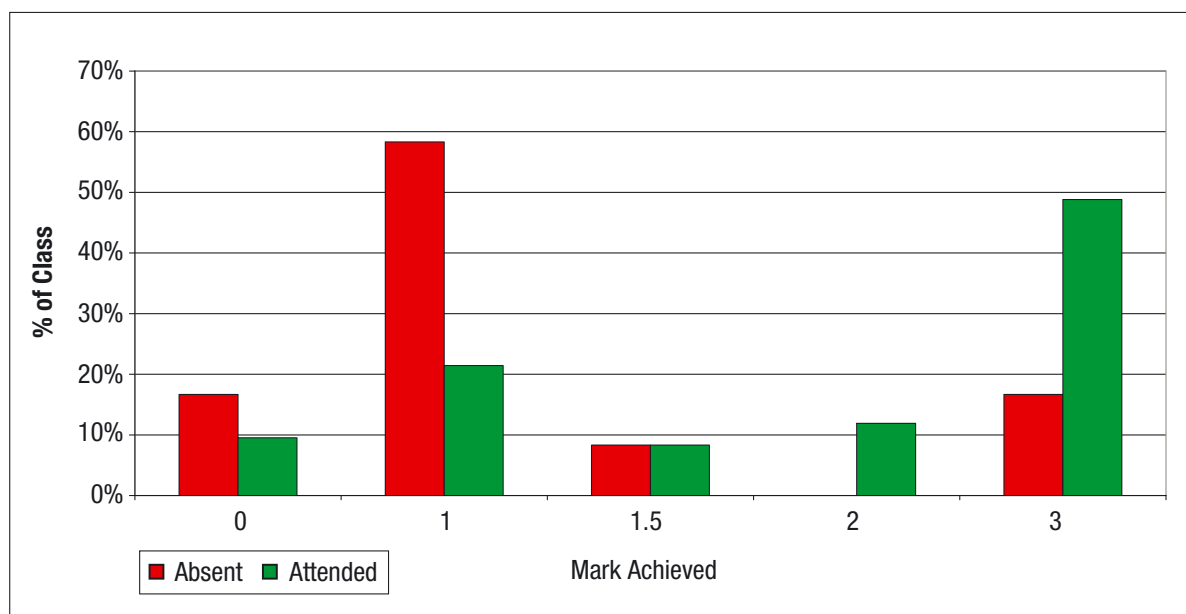
What is the pH of a solution prepared by mixing 600 ml of 0.75 M  $\text{KH}_2\text{PO}_4$  with 400 ml of 0.12 M  $\text{K}_2\text{HPO}_4$ ?

The above evidence demonstrates that on-time-feedback was successful in a number of ways. It helped ascertaining which concepts the students grasped, what constructs they need support on, and what level of procedural knowledge the majority of students in the class are able to grasp. The teaching interventions implemented during the lecture in response to the diagnostic assessment enabled more than 70% of the class to access what was made available.

Interesting findings were also obtained from a comparison of results of a question assessing the ability to apply the Henderson–Hasselbalch equation in a summative assessment. When this group of students was compared with a control group of students (the group of students that did not

attend the lecture when diagnostic assessment was implemented), it was found that almost 50% of those who attended obtained full marks, a further 12% obtained two-thirds of the marks and 9% obtained half the marks. Whilst 58% of the control group were able to *state* the Henderson–Hasselbalch equation (declarative knowledge), only 18% were able to obtain full marks for the question. This was despite all students having been required to use the Henderson–Hasselbalch equation during a practical session at which attendance was compulsory. Results are shown in Figure 5.2.

**Figure 5.2: Comparison of results of summative assessment on application of the Henderson–Hasselbalch equation between students who had attended the lecture and those who had been absent.**



It could be concluded from these findings that students appeared to have benefited from attending this lecture and that at least half of those that received on-time-feedback were able to apply the Henderson–Hasselbalch equation in a similar question in a summative assessment. Also compared were test results of the 2007 students with results of the 2006 cohort, reflecting the conceptual and procedural knowledge constructs building up to the Henderson–Hasselbalch equation. It was found that the class average for the equivalent summative assessment in 2006 was 47%, in comparison with the 62% class average of the experimental class.

Taking into account the hierarchical nature of knowledge in this field, the initial pilot study was extended to investigate whether lecture attendance affected performance as measured in the summative assessment. Specifically, we investigated whether a distinction could be made in the summative assessment performance between students who had attended all the lectures up to the point of assessment of the Henderson–Hasselbalch knowledge construct and those who had partial (up to 60%) attendance. The questions in this part of the investigation included declarative, procedural and schematic knowledge constructs on calculation of concentrations, amounts, converting different concentrations and amounts, calculation of pH, as well as the application of the Henderson–Hasselbalch equation. The results were compared between the two groups of students. The results show that there was a significant difference between the two groups. The statistical analysis is shown below:

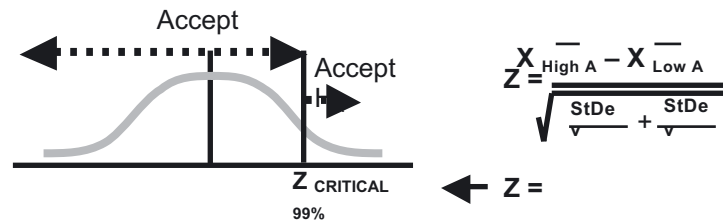
The question asked was: **Does performance in summative assessment increase with full lecture attendance?**

**Null Hypothesis (H<sub>0</sub>):**

Performance of High Attendance Students  $\leq$  Performance of Low Attendance Students

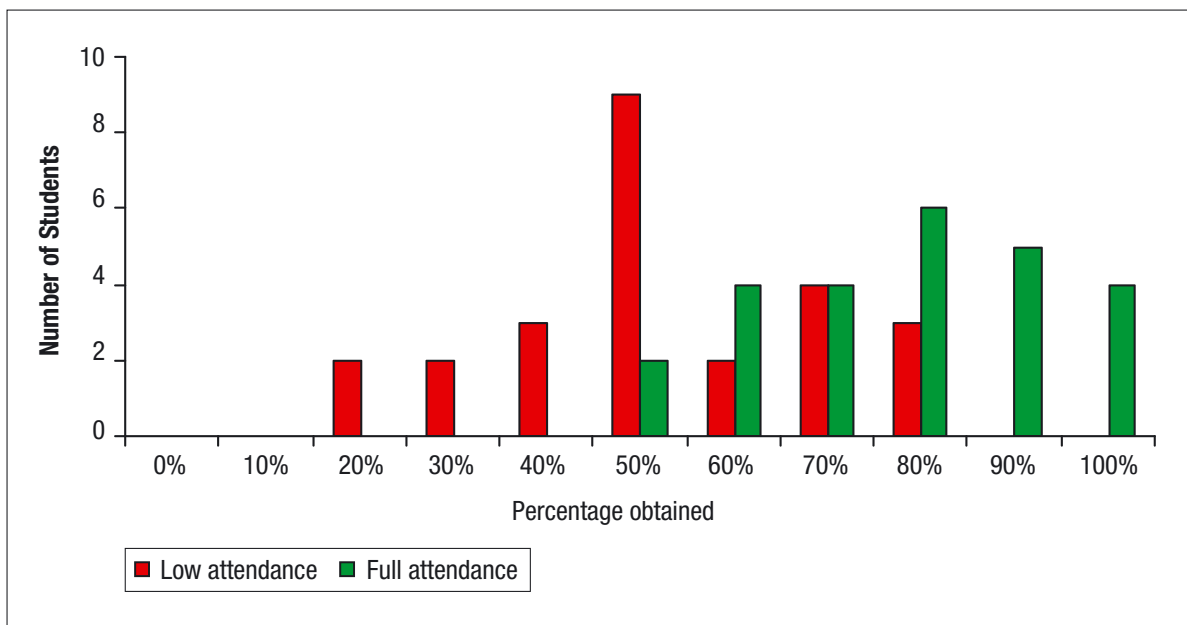
**Alternate Hypothesis (H<sub>A</sub>):**

Performance of High Attendance Students  $>$  Performance of Low Attendance Students



In addition, test results of students from each group were plotted on a graph, shown in Figure 5.3.

Figure 5.3: A comparison of summative assessment mark distribution in a group of students who had full lecture attendance with a group in which students had attended up to 60% of lectures leading to a specific procedural knowledge construct.



While further studies are required to state categorically that the observed results indicate that lecture attendance is the deciding factor influencing the performances of these two groups of students, an interesting finding is reflected in the modalities of the performances of the two groups. Most of the students in the group that attended all the lectures obtained above 80% and no-one obtained below a passing grade of 50%. On the other hand, most of the students in the group that attended up to 60% of the lectures performed at the 50% level. It appears, therefore, that lecture attendance *does* affect performance. This supports the findings comparing the results of the students who attended the lecture on the application of the Henderson–Hasselbalch equation with those who did not.

Students were canvassed for their opinions to ascertain how they felt about on-time-feedback. Opinions were also obtained from students repeating the course, as they had been exposed to the same course content taught in consecutive years in different ways by the same lecturer.

The analysis of findings suggests an observed change in the learning environment: In this regard, students indicated that the initial anonymity of the system made them feel comfortable answering questions posed to the class. Many commented that the system forced them to engage with the material as it is taught and to attempt to answer all the questions. Most also indicated that they liked the immediate feedback as they were able to self evaluate whether they had grasped the concepts taught and that they were appreciative of the lecturer's responsiveness to the feedback. Quotes from a selection of students emphasising the most salient points are given below.

I like the fact that the clickers are used in class. There is more interaction within the class. The questions also help to keep you concentrated and keep up with the class work. I also like the fact that you can tell if the class understands the work by asking a question. In that way using clickers help us to make sure if we understand the work or not. They're also just fun to use. (Student 3)

This interactive system is very good in case that you hardly fall behind with work because you listen attentively in class. The system makes you listen a lot more in class, which makes a big difference when studying because half the work is already done. (Student 8)

Students appeared to appreciate the benefits of formative assessment practices:

Clickers are fantastic. Clickers offer all of us a chance to really evaluate how much we know during lectures. It encourages attentiveness and participation in lectures and allows you to target problems in lectures. It's also a good evaluation for the lecturer since he/she can tell whether or not the class is on track by asking questions which can be answered in class with on the spot results and answers of all members. Clickers are especially useful in big classes since the lecturer is able to make an estimation of whether or not the class understands concepts presented to them. (Student 6)

They also allow students to see where the class as a whole is struggling and this could, in turn, make them feel less isolated and stressed about gaps in their own knowledge. I feel clickers are a great asset to the learning process. They encourage concentration in class and allow students to identify weak points in their knowledge. (Student 10)

The introduction of the PRS clickers has improved my understanding of the subject, as they test my ability to apply new material on the day, meaning it stays longer in my head, as I grasped the underlying concept immediately instead of having to go and learn it myself later. Also to see how I am coping in comparison to the rest of the class. (Student 12)

Apart from the instant feedback provided in class, the *Interwrite PRS* system also allows for more detailed individual feedback. Lecturers are able to track previous performance in the classroom in a subsequent analysis and to generate individual reports. In this course, this was done at regular intervals and the reports were given to the individuals during practical periods during a 'one-on-one' discussion between the lecturer and the student. These reports allowed both parties to identify the types of knowledge constructs that needed to be worked on. As students commented:

This is also good because we get given the progress sheet. This sheet told me where I needed to base and focuses my attention on when studying. (Student 8)

It was great to have a kind of instant evaluation in class. The feedback sheets were also useful as it allowed (me) to identify areas I was struggling in. (Student 9)

At first they seemed like torture to be using that small thing with buttons and it seemed pointless and useless. Oh, and it really seemed like torture because it was like I was being forced to come

to classes and could not have bunking days, like blue Mondays. But, hey, guess what, I'm a new person with a new changed mind because of the clickers. They made me not want to miss class because they take the register and made me concentrate more and understand more in class. I think they are very encouragive [sic] and supportive to students, but definitely not liked by the bunkers who like chilling at the Matrix or in the sun because for them it's pointless. They increase interaction between lecturer and students. (Student 15)

Some students realised that learning is *their* responsibility.

They provide good interaction with fellow students as well as the lecturer. They test your knowledge and how much attention you pay in lectures. Thus you know if more is needed from you. (Student 16)

Repeating students provided invaluable insight as they were able to compare this learning experience with their previous classroom experience in the same course. Attempts by the lecturer to engage these students by requiring them to complete calculations in class had been unsuccessful. The students admitted that they had previously been inclined to wait for the lecturer to show them the solution after which they had considered themselves competent to solve future problems on their own. Summative assessments had, however, soon shown them that this was not the case. Attempts to engage them in class discussions have not been successful, as, even when prompted to discuss issues in groups, certain students disengaged and left the learning to their more interested peers. D, a repeating student, had the following comments:

I think it is good learning method, because it encourages good communication between the lecturer and students. Also make us students participate in class with the lecturer and with other fellow students. I think for me it has worked because I improved my marks, and when you write a test or exam you still remember some of the questions you were asked in class. I think they should introduce it to other subjects (courses) because it helps a person understand better. Because we used the clickers you see the errors you were making in answering questions than being there in class, sitting and understanding nothing. I was one of those students who (previously) agreed on anything when the lecturer said 'did we understand or not'. But now, if I don't understand something, we do it as a group in class, and if we still don't understand, we ask the lecturer. It makes us interact more with our fellow classmates.

Interestingly, almost all the students interviewed commented on the influence of peer group learning in the classroom. Once they had keyed in their answers, students tended to compare their responses with others sitting near them and to discuss how they'd arrived at their answers while they were waiting for the rest of the class to respond. The whole class therefore appears to have now become a successful 'community of practice'. The stronger students also indicated that the immediate feedback increased their 'tolerance' of weak learners. These, in turn, stated that they felt that this system was non-threatening and that they therefore felt confident enough to participate in lectures and to be part of the combined learning process. As their responses were anonymous, they no longer felt intimidated by their more able peers. It appears though that the stronger and weaker students interact differently in class. Stronger students tend to key in their answers individually before discussing them with their peers. Weak students tend to discuss possible solutions with their peers before they key in their answers. One student pointed out that this practice might affect the validity of the judgments made by the lecturer.

The fact that you can interact and say what you thought is a bonus. Just one recommendation: As most of the students work in groups the result shown isn't always correct. Out of a group of five, one person may be right and the other four just put down that one person's answer without understanding.

That being said, at least the students who rely on a peer to arrive at an answer can identify when they do not access the knowledge. Furthermore, the on-time-feedback places them in a position to remedy their situation, either by approaching the lecturer individually, or in consultation with their more able peer.

On-time-feedback afforded the lecturer the opportunity to make valid judgements about her students' learning. It must be noted that this opportunity depends on the quality of questions used to elicit students' response. The principled point here is that questions are an important tool with which a lecturer diagnoses students' gaps in essential knowledge, identifies prior learning and diagnoses misconceptions. Reflecting on question design, Beatty et al. (2006:32) make the point that every question should have a 'threefold process consisting of a content goal, a process goal, and a metacognitive goal'. Questions designed for these purposes are very different from those traditionally used in summative assessments. In addition, immediate feedback is merely a tool which enables more effective teaching and ensures the benefit of ongoing formative assessment. Its effectiveness requires that lecturers understand the conceptual structure of the subject they teach (Muller, 2007:82), and can design questions that take students on a hierarchical path of learning, step by step from the foundations of the subject to its cumulative structures.

## Conclusion

In her review of assessment from a sociocultural perspective, Gipps (1999) analyses and comments on the aspects of power and control emanating from various assessment practices. She points out that even with the best intentions to the contrary, especially in more open types of assessment, many students defer to the authority of the instructor without deeply interrogating what ideas actually mean. Using on-time-feedback enables a meaningful dialogue during the process acquisition. It also creates a significant shift in power relations in the classroom, which is linked to the distribution of students' access to the feedback. The pace of the lecture becomes student driven, and instruction is driven by feedback and task analysis. A redistribution of power between the high achievers and slower students is also observed.

In light of the above, it is concluded that the use of immediate feedback enables this form of diagnostic (formative) assessment, and provides new possibilities for pedagogical responsiveness. On-time-feedback provides insight into students' cognitive processes as the lecture progresses; it gives an indication of the knowledge constructed by students as they learn. Careful question design is, however, paramount. On-time-feedback is also successful in ascertaining the initial knowledge level of students at the start of every lecture. This aids teaching because students' assimilation of and ability to construct new knowledge depends on their ability to relate this to prior knowledge (Newmann et al., 1996). Furthermore, results have indicated that interactive lectures provide learning opportunities that support summative assessments. Lastly, careful question design enables deep learning and helps students to make connections between various content domains.

## References:

- Beatty, ID, 2004. Transforming Student Learning with Classroom Communication Systems. *Educause Center for Applied Research: Research bulletin*, 2004(3), 1–13.
- Beatty, ID & Gerace, WJ, 2002. Probing Physics Students' Conceptual Knowledge Structures Through Term Association. *American Journal of Physics*, 70(7), 750–58.
- Beatty, ID, Gerace, WJ, Leonard, WF & Dufresne, RJ, 2006. Designing Effective Questions for Classroom Response System Teaching. *American Journal of Physics*, 74(1), 31–9.

- Bernstein, B, 2000. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique* (Chapter 9). Oxford, Rowman & Littlefield Publishers.
- Biggs, JB & Collis, KF, 1982. *Evaluating the Quality of Learning: The SOLO Taxonomy* (Chapter 2). New York: Academic Press.
- Black, P & Wiliam, D, 1998. Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1).
- Black, P & Wiliam, D, 2001. Inside the Black Box: Raising Standards through Classroom Assessment. Phi Delta Kappan International. Available online: [http://pdkintl.org/kappan/k\\_v80/k9810bla.htm](http://pdkintl.org/kappan/k_v80/k9810bla.htm). Retrieved 1 July 2007.
- Broadfoot, P, 1999. Assessment and the Emergence of Modern Society, in B. Moon & P. Murphy (Eds). *Curriculum in Context*. London: Paul Chapman Publishing.
- Butler, R, 1988. Enhancing and Undermining Intrinsic Motivation: The Effects of Task-involving and Ego-involving Evaluation on Interest and Performance. *British Journal of Educational Psychology*, 58, 1–14.
- Dann, R, 2002. *Promoting Assessment as Learning: Improving the Learning Process* (Chapters 2 and 3). London: Routledge Falmer.
- Dufresne, RJ, Leonard, WJ, Gerace, WJ, 2007. Knowledge Model. A Quantitative Model for the Storage of Domain-specific Knowledge and its Implications for Problem-solving. Extracted from a booklet accompanying a workshop for high school teachers. Available online: <http://umperg.physics.umass.edu/topics/model>. Retrieved 5 June 2007.
- Gipps, C, 1999. Socio-cultural Aspects of Assessment. *Review of Research in Education*, 24, 355–92.
- Gipps, CV & Cumming, J, 2004. Assessing Literacies. Available online: <http://literacyconference.oise.utoronto.ca/papers/gipps.pdf>.
- Hargreaves, E, 2005. Assessment for Learning? Thinking Outside the (Black) Box. *Cambridge Journal of Education*, 35(2), 213–24.
- Higgins, R, Hartley, P & Skelton, A, 2002. The Conscientious Consumer: Reconsidering the Role of Assessment Feedback in Student Learning. *Studies in Higher Education*, 27(1).
- Hirsch, ML & Gabriel, SL, 1995. Feedback Strategies: Critique and Evaluation of Oral and Written Assignments. *Journal for Accounting Education*, 13(3), 259–79.
- Killen, R, 2003. Validity in Outcomes-based Education. *Perspectives in Education*, 21(1), 1–14.
- Knight, P, 2004. Assessment of Complex Learning: The Engineering Professors' Council's New Thinking about First-cycle Engineering Degrees. *European Journal of Engineering Education*, 29(2), 183–91.
- Leahy, S, Christine, L, Thompson, M, & Wiliam, D, 2005. Assessment Minute, Day by Day. *Educational Leadership*, 18, 19–24.
- Madaus, GF, Raczec, AE & Clarke, MM, 1997. The Historical and Policy Foundations of the Assessment Movement, in A. Lin Goodwin (Ed.). *Assessment for Equity and Inclusion*. New York: Routledge.
- Maton, K, 2006. On Knowledge Structures and Knower Structures, in R. Moore, M. Arnot, J. Beck & H. Daniels (Eds). *Knowledge, Power and Educational Reform: Applying the Sociology of Basil Bernstein*. London: Routledge, 44–59.

- Morais, A & Pires, D, 2002. The WHAT and the HOW of Teaching and Learning: Going Deeper into Sociological Analysis and Intervention, in J. Muller et al. (Eds). *Reading Bernstein, Researching Bernstein*. London: Routledge Falmer, 1–25.
- Moss, PA, Girard, BJ & Haniford, LC, 2006. Validity in Educational Assessment. *Review of Research in Education*, 30, 109–62.
- Muller, J, 2007. On Splitting Hairs: Hierarchy, Knowledge and the School Curriculum, in F. Christie & J. Martin (Eds). *Language, Knowledge and Pedagogy*. London: Continuum, 65–86.
- Mumba, FK, Rollnick, M & White, M, 2002. How Wide is the Gap between High School, and First-year Chemistry at the University of the Witwatersrand? *South African Journal of Higher Education* 16(3), 148–56.
- Newmann, FM, Marks, HM, Gamoran, A, 1996. Authentic Pedagogy and Student Performance. *American Journal of Education*, 104, 280–312.
- Pellegrino, J, Chudowsky, N & Galser, R, 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.
- Rust, C, Price, M & O'Donovan, B, 2003. Improving Students' Learning by Developing their Understanding Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education*, 28(2), 147–64.
- Shavelson, R, Li, M, Ruiz-Primo, MA & Ayala, CC, 2002. Evaluating New Approaches to Assessing Learning. Keynote Address. Joint Northumbria/EARLI Assessment Conference, University of Northumbria, Newcastle.
- Shepard, LA, 2000. The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4–14.
- Slonimsky, L & Shalem, Y, 2006. Pedagogic Responsiveness for Academic Depth. *Journal of Education*, 40, 35–58.
- Taras, M, 2001. The Use of Tutor Feedback and Student Self-assessment in Summative Assessment Tasks: Towards Transparency for Students and for Tutors. *Assessment and Evaluation in Higher Education*, 26(6), 605–14.



# Reading below the surface

## Students' organisation of content and form

*Lynne Slonimsky and Yael Shalem*

### Introduction

Academic practices<sup>1</sup> are both disciplined and disciplinary activities involving specialised actions and operations, which promote the development of knowledge. The process of knowledge development is premised on conscious reflection on the ends, objects and the epistemic means of a research activity (Anderson, 1993); it involves forms of reasoning, analysis and modes of investigation that enable the critical examination of established truths, and taken-for-granted assumptions and beliefs.

In the social sciences and humanities, there is no single conception of what counts as knowledge in academic practices nor a single universal way of producing academic knowledge. However, there are some orientations, conceptual tools and operations that are common across conceptual positions. These include some partitioning of the 'form' and the 'content' of assertions (Craig, 1996), the justification of claims, engagement with established knowledge (i.e. to refute it, extend it, etc.), proof or defense of a position, principled and systematic investigation or analysis, and specialised forms of communications that can transcend temporal and spatial boundaries (Bailey, 1984).

Goal-directed actions within academic practice and the cognitive operations with which to execute them involve formalising and systematising the object of study in new ways. Vygotsky (1978) has pointed out that one of the most distinctive features of academic thinking is the ways in which concepts are brought into a relation with one another. Vygotsky refers to the kind of concepts developed through schooling and academic inquiry as 'scientific concepts' and distinguishes them from everyday concepts. He argues that the power of forms of thinking developed in academic institutions derives from the structures of thinking that are developed by systematically ordering concepts into new relations with each other. Put differently, academics impose ordering principles on the objects of investigation. They employ disciplinary questions, concepts and methods of investigation, to frame and organise the objects of investigation conceptually. From the perspective of knowledge acquisition, these knowledge practices are intended to enable inquirers to break with the naturalised logic of everyday life and to project new ways of ordering, which open up new ways of understanding reality, and new conditions of possibility for action (Muller, 2000).

In this chapter we focus on how first year students enrolled in a Bachelor of Education degree are working with the relationship between ideas. We focus on students' responses to an essay

---

<sup>1</sup> We developed this account of academic practices in Slonimsky and Shalem, 2006.

question in an exam. We are interested in how students organise ideas in their essays and what their responses can tell us about the degree of control they demonstrate in relation to academic criteria. Our aim in this investigation is to understand the epistemic means that first-year Education students adopt when investigating the object of inquiry set up by an essay question.

## Form and content in academic practice – Text-based realities

Although knowledge is the product of temporally and spatially contextualised activities, the enterprise of knowledge development hinges on a dynamic interplay between past and present, local and global contexts. What distinguishes academic knowledge par excellence is the process of de-contextualisation (Bernstein, 1990) whereby temporally and spatially contextualised knowledge is dis-embedded (Cummins, 1996) from the particular milieu or circumstance in which it is produced, and materialised in symbolic form. Wertsch (1991) argues that a key condition of possibility for the knowledge practices of the academe is engagement in what he calls ‘text-based realities,’ by which he means activity spaces that are constituted through semiotic means alone. A typical example would be knowledge in texts (academic papers published in a refereed journal) or essays.

Wertsch proposes that text-based realities have three significant properties –‘depersonalization, systemization and boundedness’, which together promote conscious reflection on the objects of study and on one’s own knowledge. We explore these briefly:

- Texts are **depersonalised** because readers and writers are generally not present to each other at the time of writing, and more often than not, are not personally familiar to each other.

Writers therefore have to anticipate a range of possible readers who do not share their circumstantial milieu by making their ideas publicly accessible to a universe of possible readers in other temporal/spatial settings. The use of linguistic, logical and conceptual means that operate at a sufficient level of abstraction and generality create a context for ideas to be communicated to readers in different temporal/spatial settings. The correlative of this is that even when texts apparently make reference to phenomena that the reader is personally familiar with, the text is simultaneously invoking a class of phenomena that extend beyond the reader’s own subjective experience. The reader thus needs to be able to distinguish between general categories or ‘types’ and particular instances or ‘tokens’. The more abstracted the ideas to be communicated, the more depersonalised they tend to be and the more specialised the language.

- Texts are **systematised** in so far as they have an internal structure and logic, comprised of parts that partially take their meaning from other parts of the text and, which taken together, constitute the sense of the text as a whole.

Such parts may be logical or linguistic, e.g. premises to conclusion, logical connectors, sentences to paragraphs, body to conclusion. In other words, they have a form over and above their content (Craig, 1996). Understanding that texts are systematised promotes sensitivity to the distinction between the form and content of knowledge and to logical argument.

- The structure and content of the text creates a **symbolically bounded semantic universe**, which simultaneously opens a world of possible interpretations of the text but delimits what is outside of the boundaries of the text.

Although particular parts of the text could be taken to mean a range of things, some of these possibilities must be ruled out because they cannot be related to other parts of the text. In other words, the sense of the text is framed through the three properties discussed above. However, the meanings of texts (i.e. interpretations) are generated through the content of texts coupled with these properties and the experiences and interpretative schemes that readers bring to the text.

It is these properties of text-based realities that promote forms of knowledge that transcend particular contexts and experiences. They also promote conditions of possibility for objectivity (i.e. sensitivity to different perspectives on phenomena), which enables one to go beyond the boundaries of one's own subjective or idiosyncratic experiences. They sensitise readers to a distinction and relation between the 'form' and 'content' of knowledge, and to the logic of argument. If students are to become full members of academic communities of practice, they must at the very least learn to work with these properties of text-based realities in both reading and in writing.

## Why first-year students? The question of underprepared students

Students, who have matriculated, are generally expected to be practised in working in text-based realities and in creating text-based realities through writing. However, a significant number of students enrolling in the humanities and social science courses in South African universities do not come to university with a sense of what it means to participate in text-based realities and are therefore underprepared for university study.<sup>2</sup>

A body of studies (Bradbury & Griesel, 1994; Steinberg & Slonimsky, 2004; Bertram, 2006, 2008; Shalem, 2004 and Boughey, 2005, 2008), as well as more informal accounts from a wide spectrum of lecturers point to some of the ways in which students who have been underprepared for university studies by their schooling tend to approach texts and epistemic practices when they first engage in university study. These pertain to the parameters of structure, depersonalisation and boundedness respectively, and include the following:

- A tendency to describe rather than to analyse, and to offer tautologies in place of justification.
- A tendency to focus on examples (tokens) rather than on principles (types), and the relation between them (this includes offering anecdotes of personal experiences in place of formulations of general principles or relating principles and particulars, or claims to alignment without explanation why).
- A tendency to write from a highly subjective viewpoint without sufficient depersonalisation, which leads to solipsistic texts and some of the other patterns discussed above.
- A failure to pull out arguments in texts or cast them (this may include syncretic lifting of isolated facts that make no sense outside of the broader structure of argument in which they are presented, poor structuring and systematisation of ideas in writing; illogical arguments and claims or discussions marked by non sequiturs).
- A tendency to include anecdotes as a justification for claims.
- A tendency to be prescriptive or normative when asked to be analytic.

At least some of these orientations to knowledge and texts are displayed in the work of any first year student. However, they tend to occur more systematically, as a *structure d'ensemble* in the work of underprepared students.

Morrow (1992) coined the term 'epistemic access'. He argued that formal access to university is not

---

2 There are many ways of explaining this phenomenon. One explanation germane to the object of inquiry in this paper is offered by Craig (1996), who distinguishes between the form and content of knowledge in academic practice. By 'form' she refers to the subtle, implicit and explicit ways of ordering and structuring knowledge that are rooted in a history of working in text-based realities that have developed over time in the history of the academy. She points out that those who have been part of this history tend to take the form of knowledge for granted, but for those who do not share this history, the rules for structuring knowledge are strange and opaque. In her paper she explores the implications of working with different permutations of form and content for promoting access to form.

enough. Students must be given epistemological access to university. By this he meant that students must be carefully initiated into those practices that enable academics to produce knowledge. As we have argued above, learning to participate in text-based realities is a key facet of the knowledge production process in universities. It could be argued that students enrolled in education courses are there to be trained as teachers, and not as researchers. However, in order to make informed and principled teaching decisions, professional teachers need to be able to move between the particular and the general, organise information and draw out implications for action. If student teachers do not learn to project beyond the ‘present and particular’, then their options as educators will be radically foreclosed.

We now turn to a case study to examine students’ constructions of content in an essay question in an exam.

## Context of the research

The course under consideration is a first year professional education course that was taught in 2007.<sup>3</sup> A team of five lecturers taught the course. Students met for three hours a week, which included lectures and tutorials. The course had a theoretical component and an applied component.

Briefly, the theoretical component focused on developing a principled account of effective teaching and the ways in which a curriculum shapes the work of teachers. It aimed to enable students to elaborate knowledge and forms of reasoning (epistemic means) that could enable them to become effective implementers of the curriculum. The applied component required students to develop principled lesson plans informed by their understanding of teaching and curricular requirements. The sample of students’ work we explore here was taken from an essay question in the June exam that tested student’s understanding of the theoretical component of the course.

Prior to the exam students wrote two essays. The first required students to reflect on qualities that effective teachers bring to bear on teaching. The second required learners to reflect on how effective teachers address different factors that may impact upon the classroom, the use of the aforementioned qualities and a range of resources or tools. Thus there was some attempt to scaffold the development of students’ thinking about effective teaching. The exam question aimed to test students understanding of effective teaching and their understanding of the official curriculum in South Africa (an outcomes-based curriculum) by asking students to explain how an outcomes-based curriculum contributes to effective teaching and learning. So the question incorporated aspects of previous essays and required students to transform this previously developed knowledge by relating it specifically to the design of the national curriculum.

## The exam question

The exam question (see Appendix 6.1) is presented in three paragraphs. In the first paragraph, students are given a preamble to the task. They are also instructed to imagine that they have been approached by parents in their community to explain to them what Outcomes-based Education (OBE) is and how learning is achieved in this type of curriculum. This is a genre stipulation. In the second paragraph of the question, students are instructed to write an essay of four to five pages explaining how OBE contributes towards effective learning and teaching. Students are also instructed that their opening paragraph should prepare their audience for what they are going to

---

3 The data for this research was collected in 2007. The course was re-conceptualised and radically restructured the following year so the curriculum that is currently being taught at the university where this research was conducted is completely different to the one under discussion in this paper.

speaking about, and the closing paragraph should sum up their presentation. The third paragraph stipulates a number of aspects students should focus on in their essay (stipulated focal aspects). These include: *what OBE is and how it impacts on teaching and learning; how the different features of the outcomes inform teaching and learning; what roles teachers and learners play in the classroom;* and also a directive to consider *what factors contribute to effective teaching and learning*. This last instruction includes further specification of what might count as other factors (personal qualities, tools and other factors that might influence the learning environment).

It seems that the overall intention of the question is to assess students' understandings of the relation between focal aspects of OBE (features, factors and roles) and effective teaching and learning processes (*Explain how OBE contributes to effective teaching and learning*). To address the intent of the overall question, the student would need to consider what s/he has learned about effective teaching and learning and factors which contribute to it, reflect on the features of OBE and then make decisions about which of all this knowledge covered in the course is appropriate for this task. Thus students have to make selection decisions.<sup>4</sup> Once students have made these selection decisions, they have to make decisions about how to elaborate and organise this knowledge to meet the stated purposes of the essay. Thus the task is to bring two sets of ideas into a relation and to explain how and why one set of ideas shapes the other.

The nature of the task is potentially ambiguous. First, students are given two different explicit instructions. They are asked to explain (to parents) *what OBE is and how learning is achieved*. They are also asked to *write an essay explaining how OBE contributes to effective teaching and learning*. The task could be construed as explaining how OBE contributes to learning, or how OBE contributes to effective teaching and learning, or how OBE contributes to effective teaching and thus, how OBE contributes to learning. Second, with respect to genre stipulations, students are first asked to imagine that they are giving a speech to members of their own local community to explain how OBE contributes to effective teaching. (*Imagine parents ...; ... prepare the audience as to what you are going to speak about*). But students are actually required to write up what they would speak about in essay format (*Write an essay*). Hence the essay is in fact to be addressed to lecturers who have explicitly set up criteria of what counts as a well-written essay. The genre stipulations potentially position students between two communities and two sets of evaluation criteria. If students were indeed to address their own community, their speech would have to be framed in relation to community concerns and priorities (i.e. to the community's evaluative criteria of the most important and significant focal aspects of OBE). But it is the lecturers and not members of their community who will be marking this essay. Therefore, students should be sensitive to the lecturers' evaluative criteria in decisions about what to select, and how to order, develop and communicate knowledge in the essay. Third, by stipulating focal aspects, lecturers signal their intention to assess several themes that have been taught in the course:

- a. whether students have understood the different factors that contribute to *effective* teaching and learning;
- b. whether students have understood what factors could influence the learning environment and what tools teachers could use to enhance their teaching in such environments; and
- c. whether students have understood the personal qualities required of teachers.

These are aspects of teaching customarily taught in first year professional education courses and would pertain to *any* curriculum (i.e. not just an OBE curriculum). If students are to address the intent of the overall question, they would need to reframe these aspects with respect to an outcomes-based curriculum. So, in fact, the question embeds further relationships, for example, how the features of OBE enhance the possibility of effective teaching; how the features of OBE

---

<sup>4</sup> For example, although students may have been taught about qualities of teachers that contribute to effective teaching, the task is not to speak about the qualities of professional teachers as such but to explain how an outcomes-based approach makes some qualities more important or salient than others.

influence learning; and how these might influence the use of tools or the personal qualities required of teachers and learners. Given the embedded nature of all of these tasks, students could lose focus on the overall intention and treat focal aspects as discrete parts.<sup>5</sup>

In summary, the instruction to give a speech to members of one's community gives the appearance that the task operates in a familiar context. The over-specification of focal aspects coupled with an under-specification of the requisite relationship to be addressed potentially obscures what the real task is.

## Developing a coding system

As our introduction has shown, an important feature of academic thinking is that it is a composite of form and content. How a student orders and relates the contents to be discussed constitutes the meaning of the contents s/he selects to focus on. To make judgements about how a student orders and relates contents, one must be able to distinguish between different forms of students' productions. Our coding system attends to this.

To develop a coding system we first selected a representative sample of twenty essays marked by each of the five lecturers (a total of one hundred essays) from four mark categories – 70% and above; 60–69%; 50–59%; and failures (below 50%). These 100 scripts were then combined and re-sorted into piles of different mark categories. We then read and began to code randomly selected scripts from each pile. Our aim was to ensure that the coding instrument we developed would be able to comprehensively describe the form and content of what was present in students' responses across all four mark categories.

On the basis of the question and the stipulated focal aspects offered in the instruction, we established a coding instrument (see Appendix 6.2). The coding instrument focuses on two dimensions:

- 1) The stipulated parameters included in the student's response (content coverage).
- 2) The structure of reasoning evident in the response (form of reasoning).

## Dimension One: Addressing the stipulated parameters of the questions

### Genre

In our initial reading of the scripts we noticed that only some students seemed to take up the criterion to imagine they were making a speech whereas others simply proceeded as if they were writing an essay with a more classical introduction or conclusion. The speech genre can be noted when students use rhetorical flourishes (e.g. *you are probably wondering*), or when they explicitly refer to an audience (e.g. *you, or your children*) or to themselves in relation to the imaginary audience (e.g. *you have invited me here today*). In order to examine in detail whether and how students attempted to engage with a speech genre, we coded for the presence or 'rhetorical moves', 'positioning self' and 'positioning audience'.

### Focal aspects

The task requires the student to focus on the following aspects:

- Outcomes-based Education (*OBE*) – explain what OBE is;
- Learning and Teaching (*L&T*) – how OBE impacts on teaching and learning;
- Personal Qualities (*Pers Qual*) – how they contribute to effective teaching and learning;

---

<sup>5</sup> As mentioned above, many students are not first language English speakers. The dense and embedded formulation makes it difficult for even a first language student to follow the task requirements, so it must be even more problematic for second language speakers.

- Tools (*Tools*) – how they contribute to effective teaching and learning;
- Environment (*environ*) – factors that influence the learning environment;
- Role teacher (*Role T*) – what role the teachers play in the OBE classroom; and
- Role learner (*Role L*) – what roles the learners play in the OBE classroom.

We added two additional categories – *context* and *evaluation*. Although students were not given explicit directives to contextualise the OBE curriculum in their speeches, contextualising concepts is a key practice in academic discourse and the course provided a background context for the development of OBE, so we were interested to see whether and how students introduced and explained the background to the curriculum. We included the parameter of *evaluation* because the speech genre instruction positions students as advocates of OBE, but a more academic orientation would entail some critical reflection on the premises of OBE and their implications for teaching and learning. So we wanted to see whether any student would critically engage with the assumption that OBE contributes to effective teaching. We used *other* to mark any content that the student included in addition to the task's content requirement.

We used all of the above categories to identify the content coverage of the exam question. This was aimed at seeing which of the stipulated focal aspects were actually addressed by the students, whether the students addressed the context and whether they evaluated if OBE does in fact contribute to effective teaching.

## Examples

The category of examples was included in the coding because, as we noted earlier, some studies of underprepared students have suggested that there may be a tendency to substitute tokens for types or to offer anecdotal evidence in place of principled explanation. Each time a student discussed an example we coded the mode in which the example was used (as substitution for explanation, illustration, extended illustration or as a way to differentiate aspects of the concepts).

## Dimension Two: the level of reasoning evident in the response

In order to characterise the quality of the students' responses, we needed to find a way to integrate the content coverage codes (Dimension One) with the form of ordering the student used (Dimension Two). In this section we describe the steps we followed to create one coding instrument (of content and form).

*Step One:* Drawing on Biggs and Collis' SOLO taxonomy and Chan et al.'s (2002) revised version (see Appendix 6.3), we developed a modified version of the SOLO taxonomy (see Appendix 6.4).<sup>6</sup> The focus on the structure of the students' response offers a picture of how each student organises ideas they have learned. It also enables the analyst to see whether the ordering of information does in fact create new or more powerful orders of meaning from the different parts that are being ordered together.

*Step Two:* We interfaced these revised SOLO levels with the categories of coverage (see above, Dimension One) in a table (see Appendix 6.2).

*Step Three:* We coded each paragraph in the student's response as unistructural, multi-structural high or low, relational high or low or extended abstract.

---

<sup>6</sup> Biggs and Collis developed the SOLO taxonomy in order to classify the quality of students' learning outcomes in different kinds of school subjects. This hierarchical model characterises the 'structure' of students' responses by focusing on what content or concepts necessary for addressing a task are included in a student's response, how these items related together, and the extent to which the overall argument is consistent and opens or closes deeper understanding of the object of inquiry. Chan et al. (2002) added more sub-levels to enhance inter-rater reliability and optimise the instrument for higher education.

*Step Four:* We classified each paragraph according to the two dimensions of the table (Appendix 6.2). This table gave us a synopsis of coverage aligned with the SOLO levels, per category. The table also included a judgement of the overall explanation in the essay as a whole (see bottom line of the table).

*Step Five:* We produced a conceptual map of the response developed in each script we coded (see Appendix 6.5). In this we highlighted the focal aspects that the student included when trying to explain the relationship between OBE and effective teaching and learning.

*Step Six:* Initial findings: By this stage of the coding, having gone through 20 scripts, it was clear that our coding system could comprehensively describe the form and content of what was present in students' responses across all four mark categories. It was also clear that none of the 20 scripts we had looked at had overall explanations that were relational or abstract. We did find some essays that had some relational paragraphs.<sup>7</sup> We then did a purposive search through all other essays, about 65 in our sample of a hundred, to see whether we could find any essays that were overall relational or extended abstract. We did not find any. The dominant form of all the responses ranged between low and high multistructural. We interpreted this finding to mean that students across all mark categories are not able to arrange the different stipulated focal aspects in a form that can create a message about the significance of the OBE curriculum for effective teaching and learning. Instead, it is as if the reader is presented with a set of 'boxes' (series of paragraphs). In one 'box' are denotative<sup>8</sup> descriptions about OBE – *OBE aims to produce critical and productive citizens; OBE is outcomes-based; OBE starts with outcomes; OBE is learner-centred; OBE aims to promote diversity; OBE aims to produce responsible citizens*. In the second 'box' there are descriptions of aspects of teaching and learning: *Lessons must instill knowledge and skills; Learners must be assessed continuously; Learners have to take more responsibility; Teachers must arrange learners in groups*. Seldom did the students explicitly relate these boxes.<sup>9</sup> In other words, they talked generally about different kinds of tools without showing how or why OBE made some tools more or less important. It was left to the reader to establish the relation between ideas in and across different paragraphs. Few examples could be found of more elaborated explanations that brought ideas into more complex chains of connections. Students offer lists of imperatives of qualities teachers must have or of functions teachers should perform, without any relation to either OBE or effective teaching. For example, *Lessons should be enjoyable* (as if one can never learn anything from a lesson that is not enjoyable); or *A teacher must be a good example to his learners* (surely this is a property of any curriculum, not just an OBE curriculum?).

In summary, from the point of view of knowledge production, there is little evidence in students' responses to suggest that students are weaving ideas together into new relations and new forms of

---

7 The following is an example of a beginning of a relational engagement: *OBE is learner-centred. This means that the teacher must plan her lessons based on what the learner knows and must work with learners' meanings. In order to find out what the learner knows the teacher would need to assess students regularly so continuous assessment is an important part of teaching*. This claim demonstrates an attempt to relate ideas together and pull out the significance or implications of this relation. Note that a more extended explanation like this would have been coded as a relational paragraph and not a multistructural paragraph because the ideas are sequenced into an argument and related together with logical connectors. We would have classified an explanation as overall relational if students did this in the majority of paragraphs in the essay.

8 A denotative description explains the meaning of a concept in different words whereas a connotative account explains the meaning of something in terms of its significance.

9 A link between a role and OBE would point to how OBE shapes this role. For example, *the teacher has to plan based on outcome. In order to plan properly the teacher must be an expert in their subject and also must decide how to mediate the knowledge. Therefore the teacher must be a designer and a mediator of knowledge*.

understanding. There is little evidence that they have transcended their everyday understanding of teaching and OBE or that they are communicating these ideas in ways that could give other possible readers (those who do not know much about OBE or how it affects teaching) a much clearer understanding. It could be argued that students would have to be superficial in order to cover all of the stipulated aspects in only five pages. So students may have made decisions to go for coverage of content at the cost of form. Another possible explanation is that the question itself was so embedded that many students did not fully understand that the aim was to focus on the *relation* between the contents of the different ‘boxes’. A third explanation is that the task was constructed as an everyday task, and therefore students did not understand that they had to create a more powerful level of understanding. A fourth possible explanation is that students do not have access to, or at least have not yet mastered, some key epistemic means that would enable them to both communicate ideas systematically, and generate new, more powerful ways of knowing, which could open access to new and more powerful ways of thinking about teaching. We develop this last explanation in the next section (Dimension Three).

### Dimension Three: The structure of reasoning evident in segments of the responses

In the next step of our research, we tried to delve deeper into how students were trying to order knowledge. At what point are they getting trapped in multistructural responses? In order to find out more we explored how students in different mark categories were trying to ‘frame’ the stipulated parameters of the question. At the most general level, the idea of framing refers to the act of putting together, shaping or building things up. Key in writing an essay (or what Linn and Miller, 2005) refer to as a ‘complex performance’) is the production of an argument – selecting a focus, deciding on what idea or concept to elaborate in more depth, ordering ideas coherently to create an educational message (a bounded whole), and, by using certain epistemic means, conveying to the reader (implicitly and explicitly) her/his control on the form of the explanation. In order to function effectively in the realm of research,<sup>10</sup> students need to learn how to engage in valid reasoning and must be able to induce, deduce and adduce information, and generate new information.

To explore framing we draw on insights from the work of Basil Bernstein and use the concept of ‘framing’ in a more specialised way. Why work with Bernstein? Throughout this paper we have argued that academic reasoning constructs new relationships between ideas and that these relationships create new orders of meaning and new possibilities for action. Bernstein (2000:27) used the concept of ‘framing’ to refer to control over principles of communication. By this he meant that the ways in which knowledge is selected, sequenced, paced/weighted<sup>11</sup> and evaluated opens or closes different orders of reality and meaning. In order to create a message about particular contents, a writer makes decisions (tacitly or consciously) about what to discuss (selection), how to weave contents together (sequence) such that particular aspects of an extra-linguistic reality (i.e. what is being spoken about – in this case teaching under an OBE curriculum) are foregrounded over others (weighting), together with some kind of evaluation or message about the nature of reality under discussion (evaluation criteria). These evaluation criteria communicate the writer’s criteria of salience or significance of the message and construe reality as more or less open (the difference between saying *the world may be like this* and *it is like this*) and may open more or less degrees of freedom for action (*you could* vs. *you must*). Together these framing moves communicate ‘ordering principles’.

---

10 Likewise, in order to make informed and principled teaching decisions, student teachers need to be able to move between the particular and the general, organise information and draw out implications for action.

11 Bernstein worked with weighting through the concept of pacing or how much time was used to develop different aspects of a message. We use weighting here.

We modified Bernstein's framing rules<sup>12</sup> and graded the adapted version of Bernstein's language of description according to three of Biggs' SOLO levels of performance (see Appendix 6.6).

- We used 'selection' to code the students' choice of content focus – that is, students' selection of specified parameters. Here we drew from our coding of Dimension One, steps 2 and 3 above. We asked: What stipulated parameters does the student select in the segment and is the coverage bounded? Does this selection offer the student a principled basis for organising and developing the response such that it addresses stipulated parameters and develops an account of the overall relationship at stake?
- We used 'weighting' to refer to the kind of detail a student provides when describing, explaining, exemplifying. We asked: Does the type of detail or elaboration the student offers serve to highlight a problem, an unintended consequence, an implication or ramification in terms of the key relation postulated by the task between outcomes-based curriculum, and effective processes of teaching and learning school content? To do this, we distinguished the form of detail developed – does the student offer a deductive or inductive explanation, an example, a qualification, normative prescription, etc. We made these distinctions by attending to conjunctions that allow for elaboration, extension and enhancement.<sup>13</sup> (For example, *It means that ... it does not say that ... in such a way ... by ... I mean ... because*,<sup>14</sup> etc.).
- We used 'sequencing' to code the linking between ideas in order to see whether the way the student links ideas in the segment as a whole enables/constrains the formation of a coherent whole or a generalisable message. We were particularly interested in how students work with the texture that they try to create through their elaboration (detail). We asked: Does the sequence of the ideas lead to a coherent message about the relationship between OBE and effective learning and teaching? In the elaboration, does the student link ideas so that the elaboration is, indeed, significant to the message? (In other words, does it add a further insight on the epistemic object of analysis?) We attended particularly to the following kinds of epistemic moves, often used to create relations between ideas (*and, it is also, therefore, namely, in addition, however, but, in contrast, on the other hand, at first*, etc.).
- With regard to 'criteria', we used the accent of the coding to assess the extent to which the student is in control of academic criteria in the moves s/he is making. Here, we were particularly interested in the ways students emphasise or linguistically mark their evaluations of the significance and importance of some ideas over others in their construction of the relation under discussion. In other words, we focused on the ways in which language users express judgements or evaluations of various types through their choice of modals (*must, could, probably, possibly, the other interesting thing*, etc.)

## Four illustrations of students' reasoning

In this section we present four illustrations of how students try to frame their discussion. Each example is a short excerpt taken from a section of the student's essay, hereafter referred to as the 'segment'. In examining the example, we first introduce what the student is doing in the sections that precede the selected segment. We also explain why we selected the segment for illustration of students' reasoning. We focus on what the student selects to focus on in the segment, and then we examine the form of reasoning as evident in the acts of elaborating and sequencing content. Lastly, we examine the ways in which the student communicates what is more and what is less significant

---

12 See Shalem (2004) for the recontextualisation of Bernstein's framing rules to the context of assessment.

13 See Eggins, 2004:47.

14 In this category, the terms *because* and *as* function as additive conjunctions and not as signifiers of a causal relation.

for her/his explanation. (Appendix 6.7 offers a comparison of the four essays from which the examples are taken from the perspective of coverage).

## One – Categorical reasoning that fails academic criteria

The student opens the essay with a classical essay introduction signalling questions she intends to discuss, including *why we are implementing OBE and how it impacts on learning and teaching ... how the outcomes inform teaching and learning and the factors which contribute to effective teaching and learning ... the role of the teacher and learner in the classroom*. This is a reformulation of the stipulated focal parameters as questions. In order to contextualise OBE, she has added a question – why are we implementing OBE? The essay is structured as a response to each of these questions. Each question is addressed in a separate paragraph.

The student begins by situating the development of OBE in the political transformation of South Africa. She then lists some OBE features (outcomes-based, aims to develop critical thinking, focuses on learner's needs, etc). She then gets to the paragraph below, which examines factors that contribute to effective teaching. The student discusses two features, each in a short paragraph, associated with OBE – interactive learning environment and the role of outcomes.

The segment that follows below was selected to demonstrate a discrepancy between the high score the student received (the student received a distinction), and the low level multistructural form of the response. The coding shows the ways in which the student tends to formulate discrete ideas as categorical statements. The segment contains three paragraphs. Paragraph one starts with a multistructural claim – *There are many factors contributing to effective teaching* – which is elaborated with a list of 'musts' about teachers' and students' roles. Paragraph two focuses on the role of parents, with a normative claim that parents must support teachers and elaborates on a form of support. Paragraph three (which we further divided into four parts), begins with a relational claim – *Tools contribute largely for effective teaching in class terms* – but proceeds with a list of different tools, each defined in terms of its function, with very little explanation on the relation between tools and the process of learning and teaching. Each of the issues provides a point on the process of learning but does not explain their relation to OBE. It is 'up to the reader' to put them together into a claim about OBE.

The segment:

[New paragraph in the original] *There are many factors contributing to effective teaching. First of all the personal qualities and attitudes of both teachers and learners are important. The teacher needs to be responsible and carefully plan every lesson. The teacher must always put the learner first and serve them whole heartedly and patiently. The teacher must also be willing to learn. Learners need to be diligent and hard workers; they need to take responsibility for their achievement in school. They must be motivated, passionate and disciplined in the learning process.*

[New paragraph in the original] *Parents also need to support learners and teachers by regularly meeting with them, assessing the learners progress and assisting with the learning process.*

[New paragraph in the original] [I] *Tools contribute largely for effective teaching in class terms. Teaching media are effective tools in which learners respond and absorb, while questioning information. Ionic media consists of media such as photographs and video clips and recordings; all allowing learners to perceive facts. Concrete media consists of field trips or outings where learners can experience situations for themselves and apply their skills easily. Abstract media includes radio recordings. There are many different media which teachers can use for effective teaching in the classroom.*

[II] *The learning environment needs to be a safe and inviting space in which learners can learn effectively and positively. Decorating the classroom with posters, diagrams and photographs will help stimulate the learner.*

[III] *The learning environment in OBE **must be an interactive and exciting one. Therefore learners are placed in groups often, where they work together and resolve many issues and questions they have.***

[IV] *Teachers **need to continuously collect resources for effective teaching as lack of finances can greatly effect the learning environment if there is not enough resources.***

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive Rule of Sequence	Discursive Rule of Evaluation Criteria
Multistructural	Multistructural	Multistructural	Unistructural
The student intends to look at factors that influence effective learning. The first factor seems to be referring to some kind of ‘a contract between learners and teachers’. The second to the parents’ role, the third to teaching tools, the fourth to learning environment and the fifth to resources. The student states what each factor entails (description) and in the case of tools (one type of factor), she states its functionality.	<p>When the student elaborates on a selected factor, the elaboration does not move beyond a unistructural type. The student does it through a proliferation of a series of personal qualities; a descriptive statement (parents); sets of definitions (teaching media); a statement and a conclusion (interactive learning environment).</p> <p>When dealing with teaching resources, the student uses a qualification (<i>as lack of</i>), stating that teaching resources are dependent on financial resources.</p>	<p>The sequence of ideas in the paragraphs above does not lead to an overarching idea about how, taken together, the factors produce/constrain a process of learning and teaching that is led by an outcome-based curriculum. Even at a discrete level, the argument is fragmented by an ad hoc sequence. If we look at the first factor (personal qualities), on the ‘contract between teachers and learners’, the student jumps from one quality to the next and there is no logical connection between them: The student opens up with the notion of <i>personal qualities</i>. She does not define the term; instead, she introduces the idea of a responsible lesson plan (which is not explained).</p> <p>The next idea seems to return to the notion of personal qualities by looking at a teacher’s care. This is supplemented by a sentence that reminds one that teachers need to learn. The paragraph ends with a categorical list of what learners must do.</p>	By using logical connectors, such as <i>the teacher needs to be; the teacher must always; the teacher must also be; Learners need to be; they must be</i> , the student indicates a categorical representation of a view. In this kind of representation there is no logical difference between an idea (interactive environment), a practice (use of a teaching tool), or a condition of possibilities (contextual reality of a school, or teacher personality). They are all of the same order.

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive Rule of Sequence	Discursive Rule of Evaluation Criteria
Multistructural	Multistructural	Multistructural	Unistructural
		<p>The idea of the learning environment is similarly treated as a list. It is repeated three times in relation to very different issues (<i>inviting place, interactive</i> and <i>collecting resources</i>); each is stated with an imperative. The prevalence in the use of the imperative disguises what is in fact a weak connection between premise and conclusions. See third paragraph, section III.</p> <p>The discussion in the paragraph produces a sequence of ideas that can easily be reorganised. The sequence could be: parents, teachers and learning contract, tools etc. Given that the list does not lead to a generalisable idea, the order of the items of the list seems not to matter.</p>	

## Two – Apparent deepening

In this essay, the student positions the ‘audience’ in a diversified social and cultural context of post-apartheid South Africa. In the three paragraphs before the excerpted segment the student discusses OBE. Consistent with the way the preamble to the question positions OBE as bringing about change, the student uses the old/new dichotomy to advocate for the ways OBE has improved on the old. The past system is described as teacher-centred, and as a system where learners memorised knowledge that was not particularly relevant to their everyday lives. Some of the principles associated with OBE (e.g. relevance, autonomy and diversity) are discussed, again, in a manner that explains a direct shift from old to new. The marker ticks various points but does not offer any comments on the content.

We chose this segment because, although the first two parts of the segment use epistemic means that suggest explanatory relationships between ideas, a closer look suggests that the weighting of ideas is simplistic and so, in SOLO levels, move between uni- and multistructural. In the third part of the segment (III), in which the student discusses the principle of relevance, the student attempts to deepen the elaboration. He points out a potential misunderstanding about OBE by qualifying the relevance of theory and the relation to practice. The link is appropriate to the principle discussed and the student’s explanation of relevance demonstrates a multistructural level. In the fifth part of the segment (V), the student shifts between principles and specific practices without any kind of indication that these are different types of knowledge.

The segment:

[New paragraph in the original] [I] *The other interesting thing about OBE is that it help teacher to plan carefully for their lessons because they are already given the learning outcomes of the specific learning area, so a teacher can use that learning outcomes as a guideline to plan his/her content in such a way that the learning outcomes can be achieved, so even the learners must be told which outcomes are required from them or are expected.*

[II] *The other principle that OBE entailed is that it is value-oriented. It means that teachers and learners can work together to make sure that each lesson that they work on have that values from assessment standards of NCS.*

[III] *Every lesson that is done inside the classroom should be relevant to what a learner would be able to apply outside the classroom even in the future [principle]. If a learner is going to fill in credit slip in future, he/she have to first start to fill in the credit slip in the classroom [example]. It is also very important for the teacher to explain to the learner that how will the lesson that they are learning benefit them in the future, remember OBE's priority is to produce a learner who will contribute to our country academically or economically for the development of our beloved country. OBE does not say that the theory must be left behind it only states that they should be integrated with practical approach, an approach in which when the lesson is completed the learner can demonstrate what he/she has learned and she/he must be able to apply it in the real world.*

[IV] *OBE forms the core of the curriculum; as it is quite understandable that we all have our individual differences, so the teacher should know all the learners in the classroom as to how she/he will address them in such a way that they will understand what was said to be done.*

[V] *As nowadays in most of the South African schools we are together as different races, so the teacher should encourage the learners to work in groups, bearing in mind that she/he must allocate the learners to sit in different races like for instance in groups of six learners, three can be black and three can be whites. By so doing she/he would be building a group of learners who consider themselves being equal and this will be a non-discriminating society, so the learners will grow up knowing that they should not discriminate one another.*

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive Rule of Sequence	Discursive Rule of Evaluation Criteria
Multistructural	Multistructural	Multistructural	Multistructural
The student seems to discuss four 'principles' of OBE ( <i>value oriented; relevance; individual differences and non discrimination</i> ) and one 'feature' (outcomes-based lesson plan).	<p>The selected feature and each of the four principles are elaborated.</p> <p>Each of the principles are explained discretely either in relation to what is to be taught (future relevant task) or how (a way of teaching or a way of classroom management).</p> <p>The elaboration of the features moves between some elaboration and some repetition – it focuses on one parameter (<i>because they are already given the learning outcomes</i>). It alludes to specialisation (<i>in such a way</i>) but does not explain it. The elaboration of the four principles varies.</p>	<p>The sequence of ideas in the paragraph above does not construct an overarching idea and does not differentiate concepts.</p> <p>Part III on the principle of 'relevance' offers a deductive explanation: It begins with a general idea (<i>should be relevant to ... be able to apply outside ... even in the future</i>). The idea is then explained by means of an example, which is justified by reference to the general aim of OBE (producing a productive citizen). The explanation ends with a qualification (it should not be understood that OBE discards theoretical role of the theoretical).</p>	<p>By using logical connectors, such as <i>The other interesting thing about OBE; The other principle that OBE entailed is and it is also very important</i>, the student indicates an additive form of discussion. The student reproduces a categorical list of ideas (four of OBE principles).</p> <p>Although the student uses epistemic means that could signal an unintended consequence of OBE and its implications (<i>it means that; remember; OBE does not say that; OBE forms the core of the curriculum</i>), that is differentiation, in fact this elaboration does not do that.</p>

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive Rule of Sequence	Discursive Rule of Evaluation Criteria
Multistructural	Multistructural	Multistructural	Multistructural
	<p><i>Relevance</i>, by means of an example (<i>if a learner ... in the future</i>) and two qualifications (<i>Remember ... OBE does not say ... it only states</i>), which relate ideas but with little elaboration, and thus is multistructural.</p> <p><i>Non discrimination</i> appears to be elaborated, but in fact the same idea is expressed (putting learners in groups will educate learners not to discriminate).</p>		

### Three – A semblance of a different idea

The student starts off by stating that he will *critically analyse how it [OBE] impacts on teaching and learning*. The student then lists what he refers to as *certain aspects* of OBE. In the next paragraph he states that *OBE has principles that are working according to them...* Each of the principles is stated and defined. For example, the student states: *the second [principle], means teachers must put needs of the learners first. It must not [be] a teacher centred. teacher must give instructions but obey learner views because some of them can make the class proceed...*

The marker ends the essay with the comment: ‘a bit more on NCS please’.

The segment starts from the fourth paragraph in the essay. It shows a case of a generalisable idea that is masked by a very weak mode of expression and a fragmented form of explanation.

The segment:

[The following points were written as one paragraph.]

- I. On OBE. *I'm sure you all see* [speech genre] *that every and each steps put learners first that means it was introduce because in old day education was the teacher centred teacher was a leader of every step but now I has changed it look fashionable but not so it helps learners to succeed more because they attended more and more.*

*The other thing* OBE edges critical outcomes *like that learners must work and communicate effectively in group or team or community work. They must also be organised and think critically. Be able to identify certain problem by communicating with others or with a teacher.*

[New Paragraph. The following points were written as one paragraph.]

- I. In OBE there are *certain* tools that make education or teaching effective and learning *The problem is* some teacher are not really qualified they lack some skills and they think teaching was their calling but it is not.
- II. The teacher *must* be patience in order to gain respect from learner *because they are from different backgrounds. This the key tool of teaching to be effective. The environment can be regarded as tools to education*
- III. Most schools are in rural places where some resources lack that cause failure that OBE discourages.

- IV. *They lack resources like electricity qualified teachers, some even lack vocabulary. This can even discourages learners to learn so parents if you choose school look at [speech genre] the tools is it in good environment or not. OBE is there to take people from the past life which was really not good for learners to pass.*
- V. *Questioning and teaching medias are also tools because if learners are given posters and ask question about them that can make sense of what is going to be done in that lesson. It also reminds them of previous knowledge and for feedback.*

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive rule of sequence	Discursive rule of evaluation Criteria
Unistruktural	Unistruktural	Unistruktural	Multistruktural
<p>There is a crucial shift in the way ideas are selected. In the first paragraph the student tries to state two ideas (the principle of learner centredness and the feature of group work), each is described in action terms of what is to be done.</p> <p>In the second paragraph, the student begins by stating that <i>there are certain tools that make education or teaching effective</i>. In stating this, the student suggests that the framing idea needs to be further differentiated but then undercuts this framing idea with a chain of other discrete contents – teachers’ qualifications, teachers’ qualities, tools, parental role, etc.</p>	<p>The shift in the quality of selection can be seen in the way the idea of ‘tools’ is elaborated. The student adds detail to what is in fact, a list of specifics. The detail includes: a notification of context (<i>the problem is</i>) – teachers are not skilled to implement teaching tools; teaching is not their calling. The student attempts to provide elaboration on tools. First, that patience is a tool (i.e. tools should also include attitudes to learners). Second, that the notion of environment must be broadened (<i>The environment can be regarded as tools</i>). In other words, the student is trying to explain that tools are not synonymous only with teaching equipment but can be regarded in a less conventional way. Implied in this is that if this idea is accepted, then the idea of tools cannot be divorced from contexts of teaching.</p>	<p>The student is ostensibly trying to relate between environmental factors and personal qualities of teachers and tools on the one hand, and effective teaching on the other.</p> <p>The combined effects of weak expression, skipping in and out of ideas (non sequitur), and the ordering of the ideas undermines the possibility of a coherent response; it thus constrains what could be potentially, in fact, an interesting educational idea – that tools are not synonymous only with equipment and thus can be seen in a less conventional way to include ‘patience’.</p>	<p>The student is using the following epistemic means: <i>the other thing like; there are certain tools that; they must also; are also; it also; The problem is; the key tool; can be regarded; are also tools</i>. There is a qualitative difference between the first five and the last four epistemic means. The first group is typically used in a multistruktural response, forming a list of concepts, features and examples. It is additive. The second group potentially signals stronger control of academic criteria but distinctions between content are blurred (ideas are not re-ordered, marked and developed so in fact it does not).</p>

## Four – Speech genre over-determines the content

In our coding we coded the overall response in this essay as ‘apparently relational’ but in fact ‘low multistructural’.

The student begins the essay by defining what the abbreviation OBE means and what outcomes stand for (*The learners will have skills, knowledge and values*). In the next paragraph, the student states that in the new system, children will be exposed to practical activities that will be of value to the life of the community. In the next two paragraphs, the student proposes that OBE has a strong principle of equality and as such can repair the wounds of the past (e.g. exclusion and alienation of rural children).

The segment that follows below was selected to show how the use of the speech genre interferes with both the selection of content items and the form of the response or how it positions the student between two sets of evaluation criteria. In the process of developing the content, the student loses the core content (‘features’ and ‘factors’ of OBE) required by the task. The student forms the argument about equality and OBE by lining up unistructural points that ostensibly appear to develop the idea of quality but in fact are stated denotatively; each point focuses on one parameter. When the student signals further elaboration (*By learner centred I mean; In this things; By doing such things; This means there would be no*), she, in most cases, resorts to what she thinks would be most significant to her audience – her community. In other words, the features or the principle mentioned are couched within a community upliftment function. The focus of the explanation is not on what the feature means in terms of effective teaching, but on how it will contribute to the community. The marker’s comment at the end of the essay: ‘Limited discussion of question, leaving out major issues depicted in the debate’.

The segment:

[New paragraph in the original] *Now let’s move from the general to beyond the classroom to get it exactly in the details how our children will acquire this skills. **You see** [speech genre] OBE is a learner centred education system where the teachers act as a facilitator of learning and a manager for learning to be effective. **By learner centred I mean** most of the activities in the classroom will be done by the learners themselves using critical and evaluative thinking. OBE aims that learners should learn about what is happening around the world, around South Africa. This will include factor **such as** economic awareness, HIV and AIDS awareness, global warming, environmental conservation and human rights and other thing **we can name them until the sun sets** [speech genre].*

[New paragraph in the original] ***In these things** I have mentioned above learners **will not only** be thought [taught] about what are they, **they will be also** thought [taught] of how to overcome them how to solve such problems if they happen in the community. This will be beneficial for the community and South Africa as a whole.*

[New paragraph in the original] *In the environmental conservation situation learners are encouraged to use every piece of material that we at home might find it useless, they are taught skills to make this thing to be valuable resource like paper cycling, collecting cans and others. **By doing such things** our communities will be ventilated while on the other hand our economy will be stable because we don’t waste to much money to buy resources from other countries or developing new one while we have our own.*

[New paragraph in the original] *In the issue of Human Rights as OBE aims that education should be equal in everyone. **This means** there would be no discrimination in black or white schools everyone would take his/her child to the school of their choice. The education that the child would get there would be equal. The would be no discrimination in terms of race as in one school all race and gender would be equal and respecting each other at school, home or anywhere should be practiced.*

In the next three paragraphs the student states OBE aims of providing skills and treating all languages equally (second paragraph); free education for those in need (next paragraph); and returning educational skills back to the community (next paragraph). The student then concludes the essay with the following paragraph:

*In conclusion I'm glad to alert you [speech genre] that OBE aims to develop a stable foundation for South African citizens. It aims to develop citizen who are both nationally and locally active. It also aims to develop citizens who will be economically active and develop job opportunities and have entrepreneurial skills. I as a teacher would advise you [speech genre] to encourage and help your children as far as possible as you can and we as teachers will take over to the next step with the assistance from the government to have a peaceful, resourceful and democratic South Africa and also economically stable.*

Discursive Rule of Selection	Discursive Rule of Weighting	Discursive Rule of Sequence	Discursive Rule of Evaluation Criteria
Relational (community) Multistructural (OBE/ effective T&L)	Relational (community) Unistructural (OBE/ effective T&L)	Relational (community) Multistructural (OBE/ effective T&L)	Relational (community) Multistructural (OBE/ effective T&L)
The student selects two features (learner-centred teaching and the use of everyday resources); she also explores the principles of equality and human rights, with an emphasis on individual and community development. The emphasis on the contribution to individual and community development appears to create function as an organiser.	The balance of detail in this response tends to be on persuading the reader (community) that the community will benefit from the introduction of OBE (freedom, equality and development). In this regard, the student is attuned to the question's requirement to address her community on the stipulated parameters of the question.  When the student says that she will get into detail (making an apparent shift to the lecturer's community), she only really provides simple and obvious points, or what we would call jargon: <i>OBE is a learner centred; teachers act as a facilitator; activities in the classroom will be done by the learners themselves.</i>	Although, in this segment, parts of paragraphs and relations between them do not flow logically, a central message is kept throughout, which is that individuals, communities and the country will benefit from OBE. In this, the student appears to point to the significance of the parts to the whole, albeit, by weak epistemic means: closing discussion on the content item far too quickly (mentioning one parameter); jumping to conclusion without sufficient explanation; and leaving out much of the detail that forms each part.	The student is using the following epistemic markers: <i>let's move from the general ... to get it exactly in the details; by ... I mean; In these things I have mentioned above ... will not only ... will also; by doing this things.</i> The first suggests that the student knows that she has to further articulate the significance of specific focal aspects; the rest show that the student can indicate focus and emphasis. Notwithstanding, the over-determination of the content by attunement to significance to the community, together with poor expression of ideas, and colloquial speech register undercuts the control over the actual relation at stake in the task.

## Discussion of findings

In all of the scripts we read (in our initial coded sample and subsequent purposive searches) we did not see students drawing on any conceptual frames or theories to order their discussions about teaching, or to articulate what counts as effective practice, or how a curriculum orders teachers' work. In other words, meta-level concepts on teaching and curriculum were notably absent. This suggests that:

- a. concepts students were given in course readings and lectures were too descriptive and did not operate at a sufficient level of abstraction and generality to offer students conceptual framing resources;
- b. such resources were offered but students were unable to see that they offer means of ordering ideas into a bounded semantic universe; or
- c. in previous tasks markers had not paid sufficient attention to the ways in which students were framing and reasoning about concepts.
- d. Or, all of the above? We do not know. An important limitation of the research reported here is that it did not do a systematic analysis of precisely what conceptual and epistemic resources were actually made available to students in the course.<sup>15</sup>

Notwithstanding, the investigation established an important lesson: for any content to be transmitted, some form of systematic production has to be followed. Bernstein's framing rules tell us that, in producing an argument, the producer needs to master four key moves: the first is selecting the focus that will be generative for more selections. The second and the third are being loyal to her selection and, thus, making decisions at every step of the process of production on how to construct this focus, and thereby on what content aspects to invest more time and space – 'time' in terms of thinking about what a content item means and expanding it for the reader; and 'space' in terms of thinking about how best to tease out its hidden parts (its insides) and where to include it in the process of articulation of the whole (weighting and sequence). Fourth is displaying command of her selection by explaining her moves (criteria). All these steps are dependent on the producer having access to epistemic means that help her to do this construction, such as justification of claims, modes of refutation of claims or of defense of a position, etc.

Using the four framing rules to look more deeply into the students' production suggests the following concluding claim: First year students do not seem to understand that the actual move of selecting a focal aspect to be addressed and content items to develop in relation to it is, already, an organising move. The problem lies in the following: The students in these courses do not seem to have a sense of which of the stipulated focal parameters are more overarching than others, and which are more or less salient to the relation to be addressed. As we noted above, one reason for this may have been that students chose to go for comprehensive coverage of all stipulated focal aspects, at the cost of form. Let's examine this in more detail by comparing the four extended examples:

The student in Example One opted for coverage of all of the stipulated focal parameters (more content, less form). Though she appears to have a handle on the phenomena, it is not clear how she actually understands the phenomena at stake. This is because, on the whole, she offers coherent summaries on each stipulated focal aspect separately. But it is only when students try to bring the concepts learned in different parts of a course into new sets of relations (rather than just summarising) that we get a real picture of what they actually understand. By opting for coverage

---

<sup>15</sup> This would have required a systematic analysis of all course communications (handouts, form and content of every reading, form and content of all lectures, and a sample of tutorials, etc.) but we did not have sufficient time or other resources to do this.

and remaining within the ‘boxes’ of stipulated content, a student may produce what appears to be a coherent essay, but such an essay may lack any overall message of significance. The fact that markers awarded this essay a distinction suggests that they focused more on coverage than on form.

Examples Two and Three show what can happen when students do actually try to open up and explore relations between focal aspects. The apparent coherence present in the kind of summarising stance adopted in Example One falls away. It is when students try to bring concepts into a relation with each other that they start to wrestle with significance and implications and try to frame a message. To transcend the world of summaries, students must re-order their thoughts, but they can only begin to do this if they have a sense of the message they wish to develop. It is at the point, where students actually try to order ideas, that issues of logic and text-based realities arise, because students have to begin to work with possible states of affairs. In trying to order, students demonstrate that they recognise (like the student in Example Three as he wrestles with the question of tools and resources) that the world of the classroom is not always clear cut and therefore begins to develop a more differentiated view of teaching, or at least aspects of it. However, markers seemed to be more focused on coverage than on how form and coverage work to create an overall message.

We have argued elsewhere (Slonimsky & Shalem, 2006) that one can only learn how to work in text-based realities if it is experienced as being functionally necessary, and if one has access to the epistemic means and other conceptual resources necessary for learning to create text-based realities. If lecturers do not pay sufficient attention to form, they will be unable to help students who have not learned to work in text-based realities to begin to do so. The whole point of an essay is that it allows one to assess whether students can order what they have learned into new relations. If one is not interested in this, and simply wishes to see whether students have a basic understanding of what has been taught and can reproduce it, then short questions would probably be a more optimal form of testing. However, if one believes that part of the power of academia comes from learning to participate in and produce text-based realities, then one cannot afford to ignore the form of students’ responses, and one must teach students that, in the end, powerful thinking cannot be had without form.

Through our coding and subsequent purposive search we noticed that students who received marks in the 60s generally covered fewer focal aspects. Many (but not all) actually tried to bring concepts into relations (less content, more form). However, those who tried to develop an extended form of reasoning in some parts of their paper usually did not sustain it and, as we saw, often did not have sufficient control over logical connectors to actually pull off a relational argument within or across paragraphs. Our argument is that this is partially a result of students not understanding the function of selection. Students do not have a sense that the aim of an essay is not simply to create ‘boxes’ of descriptions, but rather to give readers a message about these contents in a manner that shows the reader the significance and implications of these contents.

The student in Example Four is one of the few scripts we read that did find a focal point of selection. At every point in the response, the student was trying to show the relevance of OBE to the members of her community. This is evident in the ways the student framed her essay – she chose what she thought would most address key concerns in her community. This starting point allowed her to organise other parts of the essay in relation to community concerns. The essay is a relational essay with respect to OBE and the community. The problem is that, in doing so, the student did not focus on teaching at all.<sup>16</sup> From the lecturers’ perspectives, the student did not address the question and therefore failed. This is a salutary lesson in the importance of clear and unambiguous questions.

---

16 This points to one of the potential pitfalls of trying to make ‘authentic’ tasks. A community would be interested in how what is going on impacts on their lives and those of their children and are not usually interested in the minutia of a teacher’s work. If one wishes students to give a speech on how OBE impacts on teaching, it would be better to ask them to give a speech on this topic to other teachers.

## Conclusion

In our investigation we have tried to understand what epistemic means the students are using to develop the object of inquiry set up by the essay question. We argued that key to academic practice is constructing relations between ideas and in this way transforming empirical objects into objects of reflection. In constructing an essay, a writer needs to create, by means of what we have referred to as epistemic means, a bounded semantic universe. The writer needs to convey to readers of different temporal/spatial settings what to focus on and how, what is more important and what is less important, what belongs with what and what should be seen separately, what is open for further interpretation and what is a social fact, and what s/he agrees with and what s/he disagrees with. In short, the writer has to convey the internal logic of her/his construction, in particular, the relationship between the parts and the whole. Bernstein's four rules of framing help operationalise the above abstracted description of academic construction. They tell us that, when writing an essay, one is, in fact, producing a structure or a frame.<sup>17</sup> It is the production of new orders of knowing which lies at the heart of academic practice, and which should form the core of socialisation of first year students.

## References

- Anderson, CW, 1993. *Prescribing the Life of Mind: An Essay on the Purpose of the University, the Aims of Liberal Education, the Competence of Citizens and the Cultivation of Practical Reason*. Madison: University of Wisconsin Press.
- Bailey, C, 1984. *Beyond the Present and the Particular: A Theory of Liberal Education*. London: Routledge & Kegan Paul.
- Bernstein, B, 1990. *The Structuring of Pedagogic Practice*. London: Routledge.
- Bertram, C, 2006. *Exploring the Boundaries between Everyday and History Knowledge in Grade 10 Classrooms*. Paper Presented at the Kenton Education Association/Education Association of SA, Wildern.
- Bertram, C, 2008. Doing History? Assessment in History Classrooms at a Time of Curriculum Reform. *Journal of Education*, 45, 155–78.
- Biggs, JB & Collis, KF, 1982. *Evaluating the Quality of Learning: The SOLO Taxonomy* (Chapter 2). New York: Academic Press.
- Boughey, C, 2005. Epistemological Access to the University: An Alternative Perspective. *South African Journal of Higher Education*, 19(33), 230–42.
- Boughey, C, 2008. Texts, Practices and Student Learning: A View from the South. *International Journal of Educational Research*, 47, 192–99.
- Bradbury, J & Griesel, H, 1994. Text as Impetus for Learning. *State of the Art in Higher Education, SAARDHE*. Pretoria: Unisa, 323–31.
- Chan, CC, Tsui, MS & Chan, MYC, 2002. Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Students' Learning Outcomes: An Empirical Study. *Assessment & Evaluation in Higher Education*, 27(6), 512–27.
- Craig AP, 1996. Education for All. *South African Journal of Higher Education*, 10(2), 47–55.

---

<sup>17</sup> To say that a writer is a knowledge producer does not specify whether one is producing new knowledge or reproducing what other scholars have produced before.

- Cummins, J, 1996. *Negotiating Identities: Education for Empowerment in a Diverse Society*. Ontario: California Association for Bilingual Education.
- Eggins, S, 2004. *An Introduction to Systemic Functional Linguistics*. New York: Continuum.
- Linn, RL & Miller, MD, 2005. *Measurement and Assessment in Teaching* (Chapter 6). Pearson Education International.
- Morrow, W, 1992. Epistemological Access in University, *AD Issues*, 1, 3–5.
- Muller, J, 2000. Intimations of Boundlessness, in J. Muller (Ed.). *Reclaiming Knowledge: Social Theory, Curriculum and Educational Policy*. London: Routledge Falmer.
- Shalem, Y, 2004. Sign, Frame and Significance: Studying Student Teachers' Reading of the Particular. *Journal of Education*, (32), 49–80.
- Singh, P, 2000. Quality Assurance in Higher Education. *South African Journal of Higher Education*, 14(2), 5–7.
- Slonimsky, L & Shalem, Y, 2006. Pedagogic Responsiveness for Academic Depth. *Journal of Education*, 40, 35–58.
- Steinberg, C & Slonimsky, L, 2004. Pedagogical Responsiveness to Learning. Students Confronting an Unfamiliar Text-based Reality. In H. Griesel (Ed.). *Curriculum Responsiveness: Case Studies in Higher Education*. Pretoria: South African Universities Vice-Chancellors Association, 105–34.
- Vygotsky, LS, 1978. *Mind in Society: The Development of the Higher Psychological Processes*. Cambridge, Mass: Harvard University Press.
- Wertsch, J, 1991. Sociocultural Setting and the Zone of Proximal Development: The Problem of Text-based Realities, in L. Tolchinsky-Landsmann (Ed.). *Culture, Schooling and Psychological Development*. Norwood, N.J: Ablex, 71–86.

# Appendices

## Appendix 6.1: The task

The National Curriculum Statement (NCS) focuses on moving from traditional aim-and-objectives approach to an outcomes-based approach. **In Outcomes-based Education (OBE) the end results (the outcomes) and the process of learning are important, as well as the content in all learning areas.** The main features that guide teaching and learning include Critical Outcomes, Learning Outcomes and Assessment Standards.

Imagine parents of learners in your community have approached you to explain what OBE is about and how learning is achieved.

Write an essay of about 750–1000 words (four to five pages in your answer book) explaining how OBE contributes towards effective teaching and learning.

- The opening paragraph should prepare the audience as to what you are going to speak about and the closing paragraph should sum up your presentation.

Your essay should focus on the following aspects:

- **What OBE** is and how it impacts on teaching and learning
- How the different features of the outcomes in OBE inform teaching and learning
- What factors contribute to effective teaching and learning (consider personal qualities, tools and factors which influence the learning environment)
- What roles the teachers and learners play in the OBE classroom

## Appendix 6.2: Coding instrument

Genre		
	Yes	No
Introduction – prepares audience		
Conclusion – sums up		
Positioning audience		
Rhetorical move		
Self positioning		

Focal aspects coded for level of reasoning						
Focal aspect	Unistructural	Multistructural low	Multistructural high	Relational low	Relational high	Extended abstract
Context						
OBE						
L&T						
Personal Qualities						
Role T						
Role L						
Environ						
Tools						
Other						
Evaluation						
Overall expectations						

Examples					
	None	Substitute for explanation	Illustration	Extended illustration	Concept difficulty

## Appendix 6.3: Chan et al.'s revised Biggs criteria

Sub Levels	Description	Coding
Prestructural	Doesn't understand the question. Response is confused. There is no relation between the 'cue' and the answer or the cue and the response are fused into a <i>non differentiated</i> unit	0
Unistructural	One relevant datum is related to the cue. At least <i>one logical operation</i> suggests a link between the cue and the response. There is a beginning of a relation between the 'cue' and the response and that makes the one datum relevant. It is relevant to the conclusion	1
Quantitative Break		
Multistructural (Low)	<i>Two or more aspects related</i> to the 'cue' are dealt with but without elaboration; 'low' means that the several points, aspects, datum are not related and are not elaborated	2
Multistructural (Moderate)	Quite a number of items are picked up but without further elaboration. The number of ideas grows but the form of the response has not changed; it is still a kind of list of issues	3
Multistructural (High)	Many relevant points are raised and elaborated through illustration. Note that the elaboration through illustration suggests a relation (between an idea and an example) and thus Chan et al. put this on a high level	4
Qualitative Break		
Relational (Low)	Beginning of induction form: Concepts are interrelated. All the relevant data and the interrelationship Generalise ideas in some parts. The response gives a sense of the overall concepts/principle, including previous data which is selected to build the answer	5
Relational (Moderate)	Generalise ideas in many parts of the essay	6
Relational (High)	Overall generalisation of major concepts in the whole essay; <i>changes from induction to deduction logic</i>	7
Extended Abstract	Consistently generalise ideas all over the essay, and question or criticise conventional practices and /or underlying principles of the discipline.  Abstract principle Deduction from the principle Analogy Boundary of principle/deduction	8

## Appendix 6.4: A modified version of the SOLO taxonomy

Sub levels	Description	Reasoning
Unistructural	Tautological – restates phrases without actually explaining. Uses slogans without explanation or rephrasing. Denotative definition or focus on one parameter. <i>Ostensibly</i> developing idea but actually rephrasing the same idea without adding anything	Logic is syncretic – incoherent coherence OR thinking in <i>complexes</i> , rather than concepts i.e. associative chains of thinking rather than relations between categories or experiential relations (thus may use examples as substitutes for explanation)
Multistructural/low	Denotative definitions Introduces two or more points, ideas or parameters without significant elaboration of any idea. The points are not explicitly related to each other which results in an and ... and type presentation	Shows evidence of induction in development of each idea but does not relate them OR transduction – an attempt to relate ideas but makes a jump which leads to an illogical move or non sequitur OR sequencing of argument problematic, which undermines the relating logic
Multistructural/high	Several relevant points are raised and elaborated through illustration or with supporting ideas	Points are related to each other by grouping under a common category or heading. Logical development of ideas in relation to reach point but ideas are not systematically related to each other. Leads to a sophisticated ‘and ... and’ discussion
Relational/low	Denotative and functional definitions	Concepts are interrelated through the use of relating concepts or an overall concept but little elaboration
Relational/high	Overall generalisation of major concepts	Works with both inductive and deductive logic Examples may be used to illustrate relation between ideas
Extended abstract	Consistently generalises ideas and or differentiates concepts. Makes use of analogies	Logical deduction or inference and abduction from principles. Criticises or questions conventional practices Examples may be used to illustrate relation between ideas or to sharpen distinctions and relations

## Appendix 6.5: Argument map with coding – Sketch argument with structural analysis

OBE is the democratic approach to the curriculum that is more focused on producing results that are positive at the end. **This curriculum** incorporates democratic values of social justice and fundamental human rights **so it develops** each learner to full potential. (*N.S reasoning*) – OBE makes the full meaning of each and every activity that takes place in school. **In OBE** there are outcomes that guide teachers of what must come out at the end of the day. **So** as a producer **you work effectively** and be focused **if** you have a picture in mind of your product. **In the main features** I can mention critical outcomes, learning outcomes assessment standards. In addition we find developmental outcomes whereby learners reflect and explore the variety of strategies to learn more effectively. **So learners** participate responsible at school and use skills they achieve local national and global communities. (*multi-high*)

OBE is based on the learners' development **and** teachers must play an effective role in interpreting this curriculum. **They must** make sure *that learners needs are at the centre not teacher centred learning (contrast)* **that does not** contribute to effective learning. (*reason*) All subjects and domains have their different outcomes and assessment standards. **Then**, the fact that subjects are different from one another **does not mean** (*contrast*) that the learner must be blank about other subjects. (*NS transduction – multi-low*) **On OBE** there is integration of knowledge or relevant learning connected to every day knowledge. Learners concrete knowledge are allowed in the class, in fact they form part of the lesson. (*clarification*)

**Presents table of old and new curriculum in terms of active and passive learners. A set up of oppositions:**

Rote learning – critical thinking

Content based – integration of knowledge

Text book/worksheet bound – learner centred

Teacher responsible for learning – learner responsible

Content in rigid time frame – flexible time frame (*multi high – no elaboration*)

**So** this table gives clarity on what you need to do. (*transduction*) OBE is like when you get out from your house knowing where you are going rather than getting out of your house aimlessly to reach (rich) where you are going. **OBE demands cooperation** between the teacher and learner in effective teaching and learning. **So it** means that even the environment must be an inviting environment that will capture the learners' interest. (*ex subs for exp and transduction*) For example, there must be supporting materials to make learners understand the lessons. E.g. when teaching math measurements you must come with scales to class as a teacher, not just telling learners to take out rulers. We move from abstract to concrete reasoning so that all diversity will be accommodated. (*extended illustration*)

The role of effective teacher in effective learning is to be involved with learner in the field. For example, what makes a good soccer coach is the one who sometimes plays with the players in the field to acknowledge strengths and weaknesses of the players and deal with them. So the teacher must also do the same thing in the process of learning. (*ex substitute for explanation*) What is important is that we are not saying the teacher must be friends with the learners but what we emphasise is that the teacher and learner must work hand in hand to problem solve. (*clarification*) **So after** that the assessment standards will take place as a teacher you will be free to accommodate all learners needs. (*reasoning, transduction*)

In my conclusion OBE is relevant to social values. It develops the learner to ace society. (*reasoning, transduction – conclusion not argued in essence*)

## Appendix 6.6: Bernstein's language of description graded according to three of Biggs' SOLO levels of performance:

### 1. Discursive rules of Selection

Relational Strong	Multistructural Weak	Unistructural None
The student selects a key focal aspect of OBE and other focal aspects of OBE (features, factors and roles)	The student selects some focal aspects of OBE	The student's selection of content items is ad hoc

### 2. Discursive rules of Weighting

Relational Strong	Multistructural Weak	Unistructural None
The student specialises the meanings of the focal aspects of OBE. The elaboration expands dimensions of the relationship between OBE (features, factors and roles) and effective teaching and learning	Each discrete focal aspect of OBE is elaborated with relevant points or illustrations but does not go beyond a description of the focal aspect to explain the relation between OBE and effective teaching and learning	Ostensibly developing an idea when explaining a focal point but actually rephrasing the same idea without adding anything. Denotative definition or focus on one parameter

### 3. Discursive rules of Sequence

Relational Strong	Multistructural Weak	Unistructural None
Organisation focuses on the relationship between OBE (features, factors and roles) and effective teaching and learning. The discussion leads towards an elaborated generalisable message. The explanation uses both inductive and deductive logic <sup>18</sup> and examples are used to illustrate relation between ideas. Sections are related to each other by grouping under a common category or heading	Organisation focuses on discrete focal aspects of OBE. The logical development of ideas is done in relation to each point but ideas are not systematically related to each other. This leads to a sophisticated 'and ... and' discussion. Together the discrete sections produce an apparent coherent structure in a form of a list of focal aspects of OBE	The description of each focal aspect of OBE is syncretic, each discrete idea is an associative chain of thinking. Biggs and Collins refer to it as 'Transduction': a reasoning that slips about and therefore does not hold a coherent line of argument. It would include non sequiturs, and reasoning that does not meet the requirements of valid arguments because conclusions do not follow from premises

<sup>18</sup> A student reasons inductively if s/he builds up a statement from a collection of statements. In 'deduction', the student draws inferences about particulars from properties of the whole. In a 3rd logic 'abductive reasoning', the student is able to infer or deduce information and on this basis to infer and deduce additional information and implications

#### 4. Discursive rules of criteria

Relational Strong	Multistructural Weak	Unistructural None
By using epistemic means that signal implications, importance, unintended consequences, positive or negative association between factors, the student conveys differentiation of the object of analysis (the relation between features, factors and roles in OBE and effective teaching and learning) and of the significance of particular focal aspects	By using epistemic means that signal collection of focal aspects of OBE, the student conveys low differentiation of the object of analysis. Logically, each focal aspect is treated as equivalent to the other – the student conveys low differentiation of the significance of particular focal aspects	By using epistemic means that presents focal aspect of OBE as imperatives, the student forecloses the need for differentiation in the object of analysis. Each focal aspect loses its conceptual identity – the student conveys no differentiation of the significance of particular focal aspects

## Appendix 6.7

**Table 6.1: Overview of coverage in the four essays**

Example	Mark	Genre	What is OBE	OBE/ TL	Ftor/T	Ftors/T/ OBE	Role	Roles / OBE	View T	Imps Com
1	88%	Ess	✓	✓	✓	-	✓	-	Undif	-
2	62%	Ess	✓	✓	-	-	✓	✓	Mix	✓
3	50%	Mix	✓	✓	-	✓	-	-	Mix	✓
4	44%	Sp	✓	-	-	✓	-	✓	-	✓✓

In the table above, the sign ✓ means that the student has attempted to cover this focal aspect in some form. The sign - means that the student did not cover this focal aspect.

Genre was coded as essay genre (Ess) if it did not engage with the speech stipulation at all. It was coded as mixed (Mix) if the student used rhetorical devices pertaining to a speech but structured the paper for lecturers and speech (Sp) if it was structured from the logic of the community. Columns show whether students addressed stipulated focal parameters and related them (shaded columns). For example, Essay 1 covered factors (Ftor/T) of effective teaching (tools or personal qualities) but mostly without relating them to OBE and effective teaching. The column headed ‘View T’ captures whether the student develops distinctions between effective teaching and teaching as such. Where they do not capture this distinction we have noted that their view of teaching is undifferentiated. If students are predominantly categorical but occasionally show that some aspects of teaching are not clear cut, we have coded it mixed (Mix). Finally the column headed implications for community (Imps Com) focuses on whether the student relates aspects of OBE to the community concerns.

# Contributors

## Information Systems

**Susan Benvenuti** lectures in the Information Systems Division of the School of Economic and Business Sciences at the University of the Witwatersrand, which falls under the Faculty of Commerce, Law and Management. Her major focus in the IS discipline lies in Systems Analysis and Design, a field in which changes relating to both technology and methodology are fairly frequent. This has stimulated her interest in teaching and learning, and she is particularly interested in lifelong learning and self-directed learning, as both are vital to the long-term growth and success of graduates in the IS field. She is the school representative for the faculty teaching and learning committee and is currently undertaking a Masters degree in Education in the field of Tertiary Teaching, so her role in this research project encompassed aspects relating to both the IS discipline and education.

## Biology Course

**Marissa Rollnick** is Chair of Science Education and Director of the Marang Centre for Mathematics and Science Education in the School of Education at the University of the Witwatersrand. She took up the chair in the School of Education after spending 15 years in the College of Science, a science access programme for students who did not make the normal entry requirement in the Faculty of Science at Wits, first as chemistry coordinator and subsequently as director. She is the author of more than 30 refereed publications in science education in the areas of language in science, access to tertiary education and pedagogical content knowledge.

**Elisabeth Brenner**, a biochemist, has been lecturing at the University of the Witwatersrand since the beginning of 1983. Amongst the most successful pedagogies she has implemented are writing intensive courses that use writing during and out of contact periods to promote critical thinking, and the interactive *Interwrite PRS* ('clicker') technology. Dr Brenner was the recipient of the most distinguished teacher award in the Science Faculty in 2003. Due to her ongoing interest in teaching and research in education, she has since completed an MEd in tertiary education, graduating with distinction in 2009. She currently holds the portfolio of undergraduate coordinator and chairperson of the undergraduate committee in the School of Molecular and Cell Biology. This has placed her in the ideal situation to review and research teaching and curricula across the various courses offered by her school.

**Grace Moletsane** is a Senior Tutor in the Division Science and Technology in the School of Education at the University of the Witwatersrand, where she offers Life Science academic and methodology courses to pre-service teachers. She joined the Wits School of Education after 12 years at Vista University, where she taught academic education and science methodology to student teachers. She is an assessor for the Education, Training and Development Practices Sector Education and Training Authority (ETDP Seta) and has been involved with the Gauteng Department of Education (GDE) in the evaluation of Further Education and Training (FET) Life Sciences textbooks. Her areas of interest include science education, assessment and curriculum development

## Professional Education Course

**Yael Shalem** is an associate professor in the School of Education at the University of the Witwatersrand. Her research and teaching focus on teachers' work, teacher education and development, and teaching and assessment in schooling and in higher education.

**Lynne Slonimsky** lectures in the School of Education at the University of the Witwatersrand. She is interested in the psychology and sociology of knowledge and relations between them in pedagogic practices. She has brought this to bear on issues of access in higher education, schooling, education, youth development, teacher education and literacy development. Her current research explores ordering principles in pedagogic practices and their implications for the development of knowledge.



Published by:  
Council on Higher Education,  
South Africa

1 Quintin Brand Street  
Perseus Technopark  
Brummeria  
Pretoria  
South Africa

Tel: +27 12 349 3840  
Fax: +27 12 349 3928

Website: <http://www.che.ac.za>

Date of Publication: August 2010

Material from this publication may not be  
reproduced without the CHE's permission.

© Council on Higher Education, Pretoria

Designed by Jacana Media