

Continuous Assessment: Is it worth keeping?

“From maintaining to setting standards in
General and Further Education and Training”

May 2009

A joint Umalusi and Centre for Education Policy and Development Series



Improving Public Schooling Seminars

Continuous Assessment: Is it worth keeping?

“From maintaining to setting standards in
General and Further Education and Training”

COPYRIGHT 2009 UMALUSI COUNCIL FOR
QUALITY ASSURANCE IN GENERAL AND FURTHER
EDUCATION AND TRAINING. ALL RIGHTS RESERVED.

Opening and welcome

Dr Heidi Bolton

Dr Heidi Bolton from Umalusi's Statistical Information and Research Unit welcomed the audience to the seminar. She explained that this is the first of a new series of seminars that will be co-facilitated by Umalusi (Council for Quality Assurance in General and Further Education and Training in South Africa) and the Centre for Education Policy Development (CEPD). The previous series focussed on the theme 'Improving Public Schooling' and is now published in the form of booklets. Dr Bolton went on to introduce the new series of seminars explaining that Umalusi's work is focussed on maintaining and enhancing the framework of research material on the General and Further Education and Training sub-sector. The seminars planned for 2009 are about bringing this work to the public, that is, giving the public insight into what Umalusi is doing and also providing an opportunity to discuss and interact around topical and important issues in education circles.

Dr Bolton introduced the speakers starting with Mr. John Pampalis, the Director of the CEPD. He is also (non-executive) Deputy Chairperson of Umalusi. Since 1996, he has contributed to the development of new policies for basic education for post apartheid South Africa and has also worked in the areas of FET Colleges and higher education. In addition, he has authored a number of publications on school reform and the involvement of the private sector in public education as well as written books for use in the teaching of South African history in schools. Before 1991 he taught in schools in South Africa, Botswana, Canada and Tanzania.

The main speaker at the seminar, Servaas van der Berg, is Professor of Economics at the University of Stellenbosch and is also the National Research Foundation Research Chair in the Economics of Social Policy. His research and publications are in the fields of economic development, poverty and income distribution as well as the economics of education and the economics of welfare. He is also a member of the Scientific Committee of SACMEQ, a body assessing performance in fifteen African education systems. He is also a member of the Umalusi Research Forum, a group of prominent researchers that meet a few times a year to advise Umalusi on its research agenda. Professor van der Berg has served on the Ministerial Committee for School Retention.

Dr Bolton indicated that Dr Sue Cohen would be the respondent after Professor van der Berg's presentation. Dr Cohen is currently the Curriculum Co-ordinator at the Oprah Winfrey Leadership Academy for Girls (OWLAG). After some initial school teaching experience, she lectured in Geography at both the Johannesburg College of Education and the College in Giyani where she was head of the Division of Social Sciences and then acted as rector for two years. From 1995 until she took up the position at OWLAG last year, her work focused on the development of assessment practices at the IEB and also research into, and support for, practice in the field of distance education at SAIDE. In this period she also worked for the Gauteng Institute for Curriculum Development (GICD) and at Mindset, developing print and digital materials for both teachers and learners. Dr Cohen has also co-authored school textbooks in the Social and Natural Sciences. Dr Bolton therefore suggested that she was the respected voice of the Social Sciences teacher, lecturer, head of a subject, rector of a college, researcher, developer of assessment practice and of curriculum and learning material, and now of a curriculum coordinator.

Introductory remarks

Mr. John Pampalis

Mr. Pampalis began by saying that the essence of the education process takes place between teacher and learner in the classroom. He added that it goes further than this; there is also the interaction between children and their parents, their recreational activity, their educational reading and so on. However, for him, the learner's interaction with their teachers and the work that they have to do is the essence of their education experience. He suggested that when this is the way one sees education, one looks at assessment in a particular way. Learners want to pass, obtain good marks and be promoted and teachers want that, too. Mr. Pampalis reflected on the system that he experienced as a child: there was an exam at the end of the year and tests during the year but the latter did not count much. He proposed that if this is the situation then one needs to question why everything depends on the final exam. He recounted how his teachers used to say this is not a good way of assessing learners but there were no other options. Mr. Pampalis therefore argued that the need for continuous assessment grew out of a concern to make the year's work count and added that this also has the potential to motivate students. The demand for continuous assessment comes from seeing education from this perspective.

However, a public school system absorbs large amounts of tax payers' money and tax payers want to know how their money is being spent. Employers and universities want to be able to measure students and determine who would be good to employ or recruit, and therefore there are mass assessment systems. The Education Department and organisations such as Umalusi are able to see assessment from the former point of view but they have other questions to ask themselves. For example, they want to know if a student who gets 80% in Giyani is equally able as a student with the same result in Durban or Cape Town. Or in other words, they want to know about the validity of assessment systems. In a diverse country where schools have different quality of resources and teachers have different levels of qualification, there is the suspicion that continuous assessment will not produce the same results. Teachers assessing students in very different circumstances are unlikely to produce uniformity. So Umalusi, for example, contracted Professor Servaas van der Berg to do studies and determine if the instruments that are being used can measure students no matter where they are. In the matric exam 25% of the mark depends on a continuous assessment mark. So, if one has a very lenient teacher and there is not the capacity to manage quality assurance at school level, then there is the possibility that students come with the same mark but with different ability levels. So this is what the debate is about: we are looking at assessment from a systemic point of view but we also need to look at educational needs, and assessment is not the same as education. Mr. Pampalis quoted the former CEO of Umalusi: "You can't fatten the pig by weighing it".

Mr. Pampalis stated that the seminar would look at assessment from an education and school based point of view, and also from a systemic point of view. He added that he was not suggesting that the one is more important than the other; any education system has to find a way to address the problems which arise. After these introductory remarks he handed over to Professor Servaas van der Berg.

Continuous assessment: is it worth keeping?

Professor Servaas van der Berg

Professor Servaas van der Berg indicated that his presentation would focus on the 'systemic'. He suggested that Dr Sue Cohen would be likely to explore assessment issues at the level of the classroom and school, thereby providing a wider perspective on continuous assessment. He also added that his presentation was based on work done with his colleague Debra Shepherd and he therefore acknowledged her important contribution to this study.

By way of introduction, Professor van der Berg referred to the results attained by South African learners in the PIRLS literacy study, in particular, their reading score. The average for all the countries participating (noting that there were not many developing countries that participated) was 500 and the standard deviation was 100. He went on to explain that one would expect a result of 200 points below the standard deviation (the mean South African score) only 1 in 40 times. This therefore suggests that South Africa is performing a great deal lower than the other participating countries. He added, however, that the gender difference—favouring girls—is the second largest in this group of more than thirty countries, the largest being twice as large and in Kuwait. Importantly, however, only 22% of South African learners reach what is referred to in the study as the 'low international benchmark'. Professor van der Berg therefore argued that the results of these studies reflect important information about the learners in South African classrooms. At Grade 5 level, 78% of South African learners are at serious risk of not learning how to read. If compared to other countries participating in the study, it is largely oil producing countries that are in the same category as South Africa. Even in Eastern Europe less than 10% of learners are very low reading achievers.

The SACMEC II Study from 2007 provides another perspective on the South African learner; SACMEC consisting of fifteen countries in southern and eastern Africa. Here participating learners were partitioned in relation to performance, or how far away they are from the mean with a specific focus on 'between school variation' and 'within-school variation'. When compared to the other participating countries, South Africa has a very high deviation from the mean as regards the variation between schools and therefore has a high degree of inequality between schools. Professor van der Berg indicated that this is not new information but this research indicates that the variation within schools is also large by SACMEC standards; not as high as in Seychelles, Mauritius or Botswana but still moderately large. He said this research therefore shows there is a very large difference between our schools and the schooling body is quite heterogeneous. He added that about a third of those who are in danger of not learning to read at the earlier levels actually do reach matric and about half of them do in fact pass. Importantly, a large proportion of learners are close to the hurdle of the pass mark. Professor van der Berg therefore argued that it is very important that assessment is accurate so that those who have performed well enough can be separated from those who have not. Seen in this light, continuous assessment becomes very important; this mark counting 25% of the final matric mark, with the one proviso that if a school's average in the continuous assessment is more than 10 points higher than their average in the exam the continuous assessment mark is adjusted downwards so that there is only a 10 percentage point difference. This means that if a learner gets 40% in the exam and 50% or above in continuous assessment, she will get at final mark of 42,5%. It is assumed that the exam mark is the accurate one but there is always some doubt about this. Learners are therefore getting some benefit from having a higher continuous assessment mark.

Professor van der Berg then asserted that it is useful to think of assessment as a signal, as earlier suggested by Pampalis. He added that Andrew Donaldson from the National Treasury, an economist who has worked in education, stated that teachers know even more about incentives than economists because they are dealing with incentives every day. Assessment acts as a signal to the learner and is also a signal to the parent; therefore the consequences of a wrong signal are very important. Firstly, if a learner continually gets 90% even though she is performing poorly, or the assessment is too easy, she is unlikely to put any effort in. On the other hand, if a learner's marks are too low she may give up. Continuous assessment is therefore a way of signalling to learners how they should behave. It also reflects progress: a learner may have done well in one part of a paper but badly in another; therefore she knows Chemistry but does not know Physics. Continuous assessment therefore also signals where a learner should make an effort. The wrong signal can be important in terms of subject choice; many learners chose Mathematics Higher Grade but had absolutely no chance of passing in the old system. It may have been that they had the ability but, given the classrooms they were in and the teachers they had, they had no chance of passing. Subject choice also influences career choice. If, for example, a learner wants to become an engineer, she needs to know if her marks in Mathematics are good enough.

Professor van der Berg emphasised that assessment ought to provide feedback on what learners know relative to what they should know. Problems arise when assessments are unaligned with required learning outcomes. He also reiterated that inflated school assessments can lead to false security, under-preparation for exams, the wrong subject choice and inappropriate career planning. He made a distinction between inflated assessment and unreliable assessments, suggesting that the latter also give poor signals to the child who should do well or who is weak. Comparing continuous assessment to exam marks helps to identify weak teacher subject knowledge and the extent of 'noisy' signalling to learners.

He then turned attention to what is required for sound assessment to take place. He quoted from Pearson Education: "An assessment is valid when it is used for the purposes for which it is designed, allowing appropriate interpretations of the results." It is also reliable if it "provides test scores that consistently measure a student's knowledge of what is being tested." Finally, it should be aligned with and "adequately reflect the standards on which it is based". Alignment is measured in terms of the gap between the average mark in a school, or what they get for continuous assessment, and what they get for the exam. If there is a large difference, there is no alignment. That is, the difficulty level of the continuous assessment is not aligned with the outcomes that are being promoted. If an assessment is reliable, the learner who does well and has ability should perform better than the learner who has weaker capability. If this does not occur there is unreliability which is separate from whether the marks are inflated or not.

Professor van der Berg referred to a graph showing the results of eighteen Biology Higher Grade candidates (2005) to show the relationship between their exam and continuous assessment marks. If there was a perfect correlation between the two, all the marks would be positioned on the main diagonal. In this particular example, however, all the marks were positioned below this line indicating that the learners' marks were inflated. He took, for example the result of one candidate who received 55% for continuous assessment and in the exam got 30%. Therefore, the further away from the diagonal, the more inflated the continuous assessment marks were relative to the exam mark. Another learner obtained 10% for the exam and 48% for continuous assessment, thus reflecting a gap of 38%. The graph also showed substantial problems with reliability. If all the continuous assessment marks had all been inflated to the same degree, they would form a line. However, some learners with the

same continuous assessment mark produced exam results that differed by more than 20%. Professor van der Berg pointed out that one particular learner received 65% for continuous assessment and achieved 10% in the exam. In this particular class the correlation between the exam and continuous assessment marks was only 0.40.

Professor van der Berg then shifted focus to the results of another class in which the correlation was 0.80, showing that the relative position of learners in the class was better. The leniency however was still great, in one case a learner got more than 65% for continuous assessment but only 10% in the exam, this sending a very problematic signal to the learner concerned. In another Biology Higher Grade class, a teacher gave all her learners very close to 40% for their continuous assessment but they performed very differently in the exam. He suggested that this seems to suggest poor subject knowledge on the part of the teacher, or a teacher who does not know how to assess.

Professor van der Berg then focussed on the degree to which the correlation between continuous assessment and exam marks is a general problem in the South African education system. He showed a graph reflecting the relationship between average school exam results and continuous assessment marks in 2005 for Mathematics Higher Grade. Some schools were close to the diagonal, thus giving roughly appropriate signals to their respective learners. The trend line revealed that the better a school's over all exam results were, the closer it was to the diagonal, or the better the correlation between continuous assessment and the exam results. In one isolated example, however, a class got an average of 80% in their continuous assessment but in the exam the class average was in the region of 30%. He proposed that this is generally a major problem in the education system. The results for Mathematics Higher Grade reflect the results of stronger and more urban schools in South Africa. When looking at the results of English Second Language more insight can potentially be gained into what is happening in rural schools. If compared to the previous graph, these results are generally closer to the diagonal, there are also some schools which can be seen to be assessing too strictly. However, a substantial number of schools are attaining 50% for continuous assessment yet their exam mark is below 30%.

Professor van der Berg went on to discuss the adjustments made to the exam mark based on a discrepancy between exams and continuous assessment. A large number of schools' exam results get adjusted or benefit from their continuous assessment leniency. Some schools obtain higher final marks than schools at the same horizontal level because their continuous assessment mark was higher. He suggested that the system is giving the message to schools that if they assess their learners leniently to the degree of 10%, they will be given an advantage in their final results. The more weighting continuous assessment has in the final mark, the more inappropriate the signal sent to schools.

In 2003 the gaps between the exam mark and continuous assessment are smallest in English First Language, Geography and History. Over time, however, the gaps are not improving. It would be expected, if there was a feedback method and if the Education Department was strict on not being lenient, that teachers would adjust accordingly but there is no improvement in leniency. For some subjects the gap has substantially increased: for Science Higher Grade the gap has increased by 14,7%, Mathematics Higher Grade by 28,5% and History Higher Grade by 36%.

Having discussed the discrepancies between exam marks and continuous assessment, Professor van der Berg then posed the question of why there is this low correlation. Alignment could be low if different aspects of subject knowledge are tested in continuous assessment and the exam. This information is important in guiding learners how to prepare for their exam.

The correlation is expected to be high because the same subject knowledge is being tested. Some of the factors affecting individual performance are favourable home background, motivation and general ability so correlation should be expected.

The evidence suggests fairly high correlations in the exam marks of seemingly unrelated subjects, such as between English Second Language and Maths Standard Grade, Physical Science, Geography and Biology. This is another reason why one would expect correlation to be high. Correlations between continuous assessment and exam marks in the same subject should be even higher. However, in a subject such as Biology Standard Grade one can more effectively predict a learner's achievement in the exam by looking at her/his achievement in English than by looking at the Biology continuous assessment mark.

As regards the issue of correlation across the different subjects, Professor van der Berg explained that in English First Language there is an encouraging high correlation and a low gap. In Standard Grade Geography there is also a low gap. There are also high correlations for Mathematics and Physical Science, unexpectedly not for Physical Science Standard Grade however. There are rather weak correlations in Biology Standard Grade and History Standard Grade. The average Biology Standard Grade learner attains a mark of 46% for continuous assessment and 20 percentage points less in the exam. Professor van der Berg suggested that this kind of analysis could be done at provincial, district and school level to identify more specific problems with assessment.

As the exam marks of a school improve the correlation tends to improve, that is, assessment is less lenient and the signals given to learners about their ability is more accurate. The big problems occur in Mpumalanga; in Physical Science Higher Grade and Biology Higher Grade; and in Quintile 1, 2 and 3. The small gaps are found in the Western Cape and Gauteng; in Geography Standard Grade and English First Language; and in Quintile 5. Low correlations are prevalent in Mpumalanga and the Northwest; in History Standard Grade, Biology Standard Grade and History Higher Grade; and in Quintiles 1, 2 and 3. The high correlations are again found in the Western Cape and Gauteng; in Mathematics Higher and Standard Grade and English First Language; and in Quintile 5. In poorer schools the correlations are generally much lower than they are on richer schools, some of the exceptions being Geography Higher Grade and English Second Language. He suggested this may be reflective of the level of detail that is handed out to teachers about assessment.

The largest proportions of schools that assess unreliably are located in Mpumalanga and KwaZulu-Natal, partly because of their population size and the number of schools in these provinces—Mathematics stands out as a particular problem. Professor van der Berg therefore contended that if one wanted to improve assessment as a whole one should start with these two provinces.

Professor van der Berg proposed that a number of important questions arise from this information. Firstly, how accurate must matriculation marks be? He suggested that even though absolute accuracy is not possible, these marks must be a reasonably fair reflection of a child's ability. Therefore a further question should be asked: how much does one want this examination mark (which is assumed to be a fair reflection) to be diluted by a mark which depends very much on the subject knowledge and assessment ability of a large group of teachers, with some incentive to be lenient and with no strict control of how they do this assessment? Also, he asked: how much leniency is acceptable? Professor van der Berg argued that one can live with 10% leniency but not with the leniency that is reflected in these research results. He also raised the question: How reliable must continuous assessment be? He indicated that he uses a 0.6% correlation coefficient as a cut-off point, but his statistician

colleagues tell him that this is completely inadequate and it should be much higher if one wants to obtain credible information from the data.

Professor van der Berg went on to question why teachers get feedback every year yet fail to learn from this. That is, they know how their learners did but there is no reaction to these results. He felt it important to blame the provincial structures for not drawing on this important source of information about how schools are progressing and how they are assessing. Although it might be said that this information is not easily available, this need not be a problem at provincial level. Yet there is no feedback to teachers. Professor van der Berg indicated that it bothers him even more that there is no narrowing of the gaps and no increase in reliability over time. Also, in this context the focus is on matric teachers who are probably better than other teachers at assessment. Finally he questioned what mechanisms can improve assessment accuracy and argued that some accountability is needed at the level of the teacher and school. He suggested that the system used in Botswana of clusters within the inspectorate might offer some direction. It is practise in this country, at primary school level, to have a common test once a term. This test is set jointly by teachers and in some of the examples he has seen it gives the weak teacher experience of interacting with other teachers and finding out about the right way of assessing. It also gives teachers some idea of the level their learners are at and generally South African teachers do not seem to be able to manage this. Although he felt that it was not necessarily his place to make suggestions about interventions, he felt these sorts of ways of using existing resources may be useful.

In conclusion, Professor van der Berg suggested that South Africa has a system which puts too much emphasis or weighting on the matric exam. However, the system also seems to be weak in a large proportion of schools in terms of their leniency and reliability of continuous assessment. The problem is not only about what happens to the matric results as a consequence of this, but what is being systematically signalled to the learner as to how to behave in order to achieve good marks and be academically successful.

Respondent: In defence of retaining school based assessment

Dr Susan Cohen

Dr Susan Cohen began by agreeing with Professor van der Berg that the lack of correlation between school based assessment marks (or continuous assessment marks) and the final examination marks was cause for serious concern. She suggested his analysis highlighted that there are weaknesses in school based assessment practices across the schooling system and that many teachers do not have a sense of the criteria for what counts as good performance. Reports by colleagues of cluster moderation meetings and of provincial portfolio moderators suggest that a large complement of teachers do not set tasks of the required standard and also do not give sufficiently helpful feedback to learners. Therefore, although many teachers go through the motions of school based assessment and fulfil the technical requirements in terms of the portfolio tasks prescribed, their learners are not adequately benefiting from the exercise. She asserted that if this is broadly true, it is not surprising that school based assessment marks are over inflated and erratically awarded. Also, it can be expected that learners with these overinflated marks will do poorly in the final examinations.

Dr Cohen questioned whether the weaknesses in the implementation of school based assessment necessarily mean that it should be abandoned. She explained that she would

propose reasons why school based assessment is, in principle, a good thing and would also offer some thoughts as to why it is not being effectively implemented. Given its potential value in the schooling system, she felt it imperative that efforts be made to support teachers in implementing it effectively and indicated that she would make a few tentative suggestions about what measures might be taken in order to do this.

The value of school based assessment

In her first claim Dr Cohen asserted that school based assessment is an important educational complement to examinations in a variety of ways. Firstly, it allows for a more representative assessment of learners' achievement. By the time learners write the National Senior Certificate (NSC) examinations they have been at school for at least 12 years. A huge amount of learning has therefore taken place across a range of dimensions. She argued that a set of two or three papers, offering probably no more than six hours of examination time per subject, cannot adequately assess the scope of this learning. Rather, each paper can do little more than sample the wide range of knowledge, skills and attitudes that learners have developed in this time. When well set it is hoped that this sample is representative of a learner's competence as a whole and that key aspects of the learning that should have taken place will be assessed. In contrast, school based assessment can assess a greater spread of a learner's work and allow teachers to build up a rich and nuanced picture of an individual's strengths, weaknesses and levels of achievement over a longer period of time.

Secondly, assessment tasks set by a teacher for a particular group of learners can be better contextualized than can tasks set for a national pool of learners from diverse backgrounds and with very different life experiences. Some writers have argued that, while school based assessment may not be as reliable as externally set examinations, it should count equally in view of its greater validity (Steinberg, C., 2004, reflecting on Black as quoted in Parker and Rennie, 1998, p 900.). Thirdly, very often a nationally standardized paper requires learners to carry out tasks in a way that they would not normally undertake them. For example, the ability to write an essay in a formal examination setting can provide excellent information about a learner's ability to marshal their thoughts and to express them coherently in a short space of time. However, in most contexts writing does not happen in this way. It is a more protracted and reflective process, with the final product improved by the incorporation of feedback from critical readers and deepened by thinking over time. When an essay writing task is well managed in the context of school based assessment, learners are given the opportunity to develop their writing before being assessed on a final product.

Fourthly, examinations cannot assess a wide range of increasingly important skills. One of these is the ability to find and process information from a wide range of sources. As the amount of information available on any topic increases and changes rapidly, there is a real need for learners to be able to access information for themselves and to evaluate, select and organize it coherently and purposefully. While in the past there was an emphasis on text-based research (mainly from books) there is now a need for learners to conduct their own primary research, to interview people and collect field data in a range of other ways. In addition, they have at their disposal the World Wide Web and the possibility of gaining information from people in many different locations via email, chat rooms, blogs and other such web-based tools. Increasingly there is an emphasis on co-operative and collaborative learning where learners share ideas, build understanding and solve problems in a networked environment and reflect on and improve their own performance in the light of critical feedback. In addition to more conventional written responses, presentation skills have

¹ Steinberg, C. (2004) Learning Guide for Classroom Assessment. University of the Witwatersrand, Johannesburg.

become significant. Learners need to show competence across all disciplines in a variety of visual and oral modes. It is very difficult to assess this range of skills in an examination setting; it is the products of learning that can be assessed in an examination rather than the processes. In an ever changing world where learners need to be flexible and creative it seems that these processes are growing in significance and should be increasingly valued in assessment practices. This is made clear through the Learning Outcomes in many subjects. The Natural Sciences provides a good example: Learning Outcome 1 in both the GET and FET Curriculum Statements is about 'doing science'. This relates to putting the scientific method of enquiry into practice: formulating hypotheses, collecting data, analyzing and synthesizing data, drawing conclusions, presenting findings and reflecting on the investigation process. Some of these aspects can be tested in examination settings through skilfully constructed questions but the actual ability to 'do science' cannot (Orpwood, 2001).

In addition, if school based assessment is coherent with the curriculum outcomes, it should strengthen learners' performance in the final examination. While some discrepancy might be expected if different skills and competences are being assessed, the message of overall competence should be similar. In all the Subject Assessment Guidelines the prescribed tasks for school based assessment overlap to some degree with the kinds of tasks that form the basis for an examination, for example, tests and internal examinations in all subjects and the range of essays in languages. Even where there are seemingly different tasks, at least some of the fundamental competences are common. The difference in marks between school based assessment tasks and the examination should consequently not be great. For instance, although there will be differences in assessment criteria for 'putting forward an argument' in oral or written form, there are also several key criteria common to both—such as having valid and current information and being able to develop a line of thinking in a coherent, persuasive and logical way. Learners who excel in one format are unlikely to perform really dismally in the other.

Dr Cohen summed up her first claim in support of school based assessment by stating that examinations can test many, but not all, important aspects of learning and so cannot offer valid assessment of the curriculum. In many areas the gaps in what is possible in examination conditions can be filled by school based assessment tasks.

Dr Cohen next asserted the claim that school based assessment is essential if assessment is to play its proper role in supporting learning. Firstly, the links between teaching, learning and assessment are being increasingly formally recognized and encouraged by many education writers. It is essential for learners to engage in a wide range of meaningful learning activities and to receive constructive feedback in order to deepen their thinking and develop important skills. They need opportunities to practice and to reflect on their progress. Teachers, too, need to engage with learners and with the products of their tasks while they are working in order to gauge 'where they are at'. The information from on-going assessment enables teachers to identify their learners' strengths and weaknesses, gain insight into their teaching practices and adjust their strategies so as to more effectively meet their learners' needs. Thus in essay writing, for example, the iterative process provides the teacher with valuable information and can act as a useful diagnostic tool.

Having argued for the value of school based assessment Dr Cohen said that the policy for including school based assessment in the compilation of the final mark at both external examination points is based on sound educational principles. In addition, she felt that including school based assessment in the final mark sends a strong message to teachers that

² Orpwood, G (2001). 'The Role of assessment in Science Curriculum Reform'. In *Assessment in Education*, Vol. 8, pp 135 -151.

their assessment of their learners is important and that their judgment of learning is valued. It should, therefore, encourage teachers to take this aspect of their work seriously. It also signals to learners that work done in the process of learning, and which is not examinable in traditional ways, has value.

National policy and guidelines are intended to support the implementation of sound school based assessment practices. In each subject they prescribe a certain minimum number of tasks that are to be assessed by teachers each term. In order to introduce some level of standardization and to encourage a diversity of assessment tasks policy also prescribes the nature of these tasks, for example, a research project, a certain type of essay or a controlled test. Policy also specifies the weighting to be allocated to each task in the compilation of the final school based assessment mark. Exemplars of tasks and of assessment tools are provided to support teachers in understanding the standard required. Moderation processes are also prescribed at schools, district cluster meetings and by nationally appointed moderation teams.

Dimensions of failure of school based assessment

Dr Cohen referred back to Professor van der Berg's evidence that the implementation of school based assessment is not achieving its potential in South African schools. On the one hand, a large complement of teachers simply does not have the expertise to competently implement the policy. She stated that many teachers have not had adequate training and suggested that there is wide consensus that C2005, in both the original and revised versions, has been implemented without due regard for the developmental needs of teachers. Training has to a large extent focused on the technicalities of the curriculum and its associated assessment practices. The recently published report of the ministerial committee on a National Evaluation and Development Unit refers to several studies in which teachers have complained about not being offered genuine and effective development support on the implementation of curriculum and assessment policies (2009:17). Large numbers of learners are therefore not benefiting from the implementation of school based assessment policies.

Dr Cohen asserted that, on the other hand, some of the official policies and guidelines stifle best practices in school based assessment in different contexts. Many teachers report that the prescriptions of the school based assessment procedures limit their ability to fully use it as a developmental tool for their learners. They find that by having to commit to a predetermined assessment plan they are unable to respond as flexibly to learners' needs or to assessment opportunities that present themselves because of events in the broader environment. There is also a sense that, where the details and topics of tasks are actually specified, they are not always appropriate for a certain group of learners. That is, they are being linked to an inappropriate context, are at the wrong level, or demand resources which are not readily available. Most subjects require an examination in both the second and third terms of Grade 12. This means that a large amount of time—in an already short year—is taken up with this component of the school based assessment plan. Concerns have been expressed about the marking grids and rubrics that make the marking and recording processes onerous, especially in large classes. There is a sense that this takes up time that might be better spent on other activities. Some teachers feel frustrated that so few tasks are considered for the school based assessment mark, noting that they do far more than what is prescribed. They would like their learners' broader achievements to be acknowledged in some way. Others feel that the pressure to get through the prescribed tasks means they do not have time to allow learners to

do additional work. They are therefore unable to improve on skills and understanding that a task shows to be below standard as they need to move on to the next task, and get it ticked off their list. Learners end up doing tasks for the sake of getting them done, rather than as an integral part of the teaching and learning process. Dr Cohen indicated that Pryor and Lubisi argue that if teachers are encouraged to conceptualize school based assessment as primarily concerned with the collection of evidence, and not as part of formative assessment, it will actually 'get in the way of it.' They suggest that:

...it is the better focused instruction that might accrue from more responsive and more frequent informal formative assessment that is liable to improve learners' chances in summative assessment. If the aspirations of the new curriculum are to be met, it is important therefore that assessment of learning is not swamped by the demands for formal continuous assessment.

(Pryor and Lubisi, 2002: 679)

Dr Cohen argued that even the teachers who have the competence to set tasks that are of the appropriate standard and give relevant and helpful feedback feel that they are going through a technicist exercise and are not able to meet their students' needs as they were previously able to do. In addition, many teachers suggest that the guidelines regulating school based assessment, such as those that require the use of externally set tasks and rubrics, have taken away meaningful responsibility and decision making in a core aspect of their professional lives.

Dr Cohen stated that she believes it is necessary to prescribe the number and type of school based assessment tasks to be included in portfolios if some sort of standardization is to be achieved. In addition, departmental guidelines should specify a range of tasks to ensure that learners are exposed to these forms of assessment tasks (for example: model making, data response questions, investigations and case studies) as well as the more conventional written responses to questions. This specification helps to support curriculum change as it encourages teachers to include these activities in their teaching repertoire. However, she asserted that while some specification of tasks required in portfolios might be necessary, the degree of specification is increasing in an attempt to support and control the work of weak teachers. Increasing specification does not, however, serve the practice of school based assessment well. Firstly, specification itself does not support teachers in designing tasks and assessing performance if they do not have the competence to do so. This is particularly true where the guidelines are difficult to follow and where exemplar tasks and assessment tools are of low quality and do not model good assessment practice. Secondly, specification can mitigate against the integration of assessment in teaching and learning. It can also dampen innovative practises by imposing time frames, tasks and tools on teachers who would otherwise have made more appropriate decisions for themselves. She argued that, to a large extent, the high degree of specification that dominates school based assessment is a bureaucratic solution to the problem of poor teaching and assessment in too many of our schools. This solution undermines the professional judgment of the many good teachers in the system and does not help weak teachers to gain a sense of what is appropriate and what is not. Dr Cohen noted, however, that removing the requirement that school based assessment forms part of the final mark will not address these concerns in a constructive way.

* Pryor, J and Lubisi, C (2002). 'Reconceptualising educational assessment in South Africa – a testing time for teachers' in *International Journal of Education Development*, Vol. 22, Issue 6, pp 673 – 686.

In schools where teachers lack the expertise or the motivation to implement effective school based assessment, taking it away is likely to cause a reversion to a stronger emphasis on infrequent paper and pencil summative assessment, less formative assessment and a narrowing of the range of tasks. Also, the role of assessment in teaching and learning will likely be diminished and teachers will have little incentive to embrace improved assessment practices. All of this will undermine the implementation of the National Curriculum Statement. Dr Cohen suggested that school based assessment, even in a diminished form, would continue even if it did not contribute to the final mark. Therefore when incompetent teachers set and grade work, marks would continue to be erratic and overinflated and the wrong signals about competence would continue to be given to learners. Also it follows that, where teachers are doing good work, the value of this would not be acknowledged nor reflected in their learners' final marks. She therefore contended that, instead of scrapping the school based assessment component of the final mark, it is necessary to find ways of strengthening it.

As regards how to strengthen school based assessment, Dr Cohen reflected that this is a difficult area. Key, however, is the need for teacher development in assessment practices. That is, teachers need help in understanding more than the technicalities of what is required, which is what they have been getting from most of the training offered thus far. Rather, they need training that helps them set appropriate tasks for their learners, align their tasks to Assessment Standards, set appropriate standards, assess learners against these, analyse learners' errors and provide formative feedback. Alongside this there is the need for a critical review of the Assessment Guidelines and exemplars to ensure that they support good practice. She also argued that the requirement of two examinations in the matric year should be reviewed. These take large chunks of time and put pressure on the teaching and assessing programme.

At present, cluster meetings are a major focus for teacher development initiatives. They are the places where peer moderation of the prescribed assessment tasks is carried out. As such, they have enormous potential for bringing about improvements in assessment practices. It is here that teachers could share ideas for the design of assessment tasks, critique each others' tasks and assessment tools, engage in meaningful moderation of learners' work, and attempt to arrive at a common understanding of standards through these activities and the discussion of exemplars. Unfortunately, teachers' reflections on cluster meetings suggest that many, for a variety of reasons, are not fulfilling their potential as a catalyst for improvement in practice. The moderation of portfolio tasks is often scantily done and based on moderating tools that tend to focus on compliance with policy and the technical aspects of marking, rather than deeper quality issues. For example: have the correct number of tasks included and is there school stamp etc? There is little genuine engagement with discrepancies in the rigour and demands of tasks, nor with concerns about ways in which tasks are formulated. In many districts certain subjects have no facilitator, and where one is appointed, he or she is not always able to give the necessary support to teachers and so acts as little more than a conduit for instructions. Dr Cohen explained that she has been at meetings confined to reading out a circular and explaining its details. Thus, if clusters are to play their envisaged role as loci of teacher development, she felt that serious attention should be paid to making them more effective. This will involve drawing up some sort of programme for developing teacher competence. It is necessary to ensure that district facilitators have adequate capacity or expertise should be engaged from elsewhere. It is equally important to critically review the current models of appointing cluster leaders, their duties, and at the entire modus operandi of cluster meetings. In addition, it is essential to conceptualise and implement alternative complementary strategies for improving teachers' assessment practices.

Dr Cohen suggested school based assessment could also be strengthened by acknowledging differences in teacher competence and exempting teachers who are

clearly managing school based assessment effectively from some of the more prescriptive guidelines. In particular, their learners should be exempt from writing common exams and from engaging with common tasks. Teachers should be encouraged to do tasks in addition to those prescribed in subjects where this is possible. They should have more flexibility in deciding which tasks best reflect their learners' achievement of the outcomes. District officials charged with the responsibility of monitoring the implementation of school based assessment should be encouraged to use their discretion in evaluating deviations from the prescribed procedures. In the same vein, there needs to be more rigorous follow up in schools (and of teachers within them) that are clearly not implementing school based assessment effectively. Although, in principle, reports of poor performance from provincial moderating panels are submitted to districts and schools these do not always result in constructive and targeted interventions.

In conclusion, Dr Cohen affirmed that the concerns raised by Professor van der Berg are important. The examination system alone, at this stage, does constitute a more efficient and more reliable mechanism for providing standardized results and for effectively ranking our learners. However, these attributes are not our only concern. Of equal importance is the improvement of the quality of teaching and learning in our schools and ensuring sound school based assessment practices is an essential component of this. Professor van der Berg showed that school based assessment practices are generally very weak. Ignoring this component of the final mark, however, sends an equally negative message about the competence of the teaching profession, increases public lack of faith in the system, de-motivates many excellent teachers and deters further aspirant entrants into it. More importantly, it will do nothing to strengthen assessment practices in schools and this is where our energy must be concentrated. In the meanwhile, she suggested the statistical moderation of school based assessment marks should continue to ameliorate the effect of school based assessment marks on learner results.

Discussion

Dr Bolton facilitated the discussion, beginning with a participant from Umalusi who made two comments. He firstly endorsed Professor van der Berg's point about the importance of feedback after the final exams and suggested this should be the responsibility of provincial education departments. Various provinces have adopted different approaches and this is an area that needs strengthening. Although this is a lot of work it must be done and some provinces are performing well. He felt feedback should be done incrementally targeting weak subjects initially, and later broadening out to include others. Teachers need to get a sense of the required standard to help them to identify their own shortcomings. He added that teacher development is a huge and extremely expensive task but provision has thus far been insufficient. This intervention could also be undertaken incrementally; subject advisors need to be targeted first so that they are well equipped to guide teachers and there needs to be more effective leadership and management of clusters.

The next speaker from the Department of Education congratulated Professor van der Berg and Dr Cohen on their excellent presentations but indicated that she thought it important to clarify that Professor van der Berg's research was focussed on the years 2002 to 2005, whereas Dr Cohen's comments were more specific to the interventions of the Department of Education subsequent to this period. She explained that the department's focus, in lieu of the problems reflected in Professor van der Berg's research, has been on making sure that the teacher is given adequate guidance. She also referred back to one of the graphs shown by Professor van der Berg which showed that a teacher had given all her students

40% and suggested that this teacher never actually assessed her learners in the first place. The Education Department's attempt to specify that a teacher must do seven tasks for History and eight tasks for Mathematics is therefore an attempt to make sure that learners in these situations actually are assessed. Turning attention to the results in 2008, she said the 26 000 learners who did not get their results on time was a reflection of a greatly reduced number of learners who never previously had school based assessment. The Education Department is therefore putting fire under teachers to complete their assessment in schools. She also indicated that it was important to take into consideration that the exams changed in the period 2002 to 2005; initially the exams became easier and then suddenly in 2005 they became harder. She finally clarified that there is no policy that prescribes that every school must undertake two exams in a year; in fact the department strongly suggests that this should not happen as this impacts on teaching time. Also teachers are expected to use their professional judgement in relation to the guidelines, and these are 'guidelines', not policy.

Another member of the audience indicated she had been an external moderator for Umalusi and found that in recent years there was a serious lack of moderation of school based assessment happening in the six provinces that she had visited. She suggested that 'broad banding' has removed middle management in schools. A Head of Department is no longer necessarily a subject specialist and therefore teachers are allocated this task, but teachers do not take kindly to their peers moderating their learners' work and the moderators are not paid for this extra work. Moderation is therefore often not completed. Sometimes moderation is monitored by the Deputy Principal or Principal but this is not always the case. The speaker described how she had seen numerous files where there was no stamp or signature, or any sign that anybody had ever looked at the work. It was therefore her observation that although there was some monitoring, there was hardly any moderation and this was a serious problem. She felt this could be addressed by some kind of compensation, promotion or acknowledgement for the work that is being done. She further stated that, in most schools, the people who are teaching English are second language speakers and their level of competency is not always what it should be. She also agreed with what Dr Cohen had said about training and its focus on technicality and that there is very little support for the development of the subject knowledge and skills required in assessment. She felt this needs to be addressed; teachers have no idea how to use rubrics to assess essays. Also, there are some clusters that are working very effectively, but in most cases they are not, especially when teachers have to travel far to get to them. However, they are also not working where there is only a short distance to travel. She felt this may be because there is a lot of work involved in organising and facilitating cluster meetings and this is also not recognised, there is also no remuneration for transport. Although clusters can be used to develop skills this is not happening. In conclusion she stated that these issues need to be looked at and addressed at grass roots level if school based assessment is to become more functional.

Another participant said that he had enjoyed the sobering reflection of 'reality' as presented by Professor van der Berg's research. He agreed that it was important to question the signals given to learners through school based assessment. That is, they are being told they are capable of between 60 to 80% yet many learners achieve about 14% in the final exam. He proposed that the cause of this exaggerated performance lay in the language of teaching and learning, and in the language in which papers are set. He suggested that the correlation between English and Biology results supported this argument. He then proposed that, in the case of the PIRLS research findings, about 80% of learners in South Africa are taught in mother-tongue and then taught in English for two years. When tested against other countries where children have been taught in English for five years, he felt it obvious they would perform more poorly. He also argued that research suggests learners perform better in Mathematics that is not word related. Therefore there is a basic problem with the manner

in which the question papers are set for external examinations. Learners do not understand the questions and in the majority of black schools in South Africa learners study English on the Lower Grade and their mother tongue on the Higher Grade. Therefore the proficiency and ability of a child is compromised in an exam. He felt the Department of Education needs to find a solution to this problem and suggested that alternative forms of assessment need to be used so that a closer correlation in results can occur and the frightening situation presented in Professor van der Berg's research averted.

Professor van der Berg responded to some of the issues raised by the audience, agreeing that feedback after the exams is crucial and that there is far too little feedback in the system. He added that this requires close communication through all the channels down to schools and therefore this process often breaks down. He suggested feedback should be simply focused on exactly how schools performed in school based assessment and how they did in the exam. The correlation or lack of correlation should be made clear and schools should also be informed about how their results compare with the national average. This would give some indication as to how the school had done, and he suggested an additional paragraph explaining how to interpret these results would be useful. Even though there are time-lags, this is what is needed. However, what finally counts in making this process meaningful is a motivated principal who cares.

Professor van der Berg also went on to say that he did not agree with the views expressed about 'language' being problematic. This is not to say that it is not an important issue; Umalusi has done research on weak performance as a result of language competency in technical subjects. However he clarified that the PIRLS research was conducted in a language chosen by participating schools. More than two thirds of these schools were in fact tested in one of the African languages and yet their results were still weak. He therefore felt it was not a question of English and added that other African countries experience the same problem and yet they were not similarly affected. There have been some attempts by the department since the PIRLS study to deal with this issue but it is a problem of such major dimension that it is not possible to see rapid results. He proposed that school based assessment should not be thrown out, however, there is enough information to suggest that the wrong signals are being sent to learners therefore the weight of school based assessment needs to be seriously considered.

Dr Cohen agreed that it is important to distinguish the time frame in which Professor van der Berg's research study had been completed from the Education Department's subsequent intervention; however, increased specification is not enough. She added that in the case of English Higher Grade, school based assessment is not noticeably different from how they assessed in the past. In Biology the difference is huge; in the past learners were required to 'spit out' information on a page but now the style of question and school based assessment requirements are very different, for example, there is a strong emphasis on doing empirical research. She argued that providing exemplars for teachers who do not know how to facilitate these processes is not enough. She further suggested as regards the issue of feedback that teachers need added support about what to do to make their results better. In her experience, teachers generally want to do better and are frustrated that they are not getting this kind of help. She referred to a meeting she attended in which mediation was limited to going through typographical errors in the assessment paper. In this context she also challenged the facilitator as to why only the final answer was rewarded but was told that this was not going to be discussed. In this context there was therefore no substantial feedback or support.

After the above responses, Dr Bolton took a few more questions and comments from the audience. A SADTU representative questioned the role of teacher training institutions and whether the way teachers are being trained at institutions meets the current needs of the classroom. He also pointed to the role of the Higher Education Quality Committee in terms of the quality assurance of various in-service training available for teachers. He added that the in-service training initiatives undertaken by the department also need to be scrutinised as to how they are empowering teachers to be better assessors. Another person added that the standard of the exam fluctuated from year to year and that this needed to be considered in relation to the lack of correlation between school based assessment and the exams.

Professor van der Berg agreed that it is very important for teachers to learn how to assess and commented that teacher training institutions need to engage with this area. He also suggested that subject knowledge was relevant. However referring to the PIRLS study, he felt that learners' performance in reading at Grade 5 level cannot be attributed to their own difficulty with reading. Rather, it is the lack of knowledge as to how to get the information across. Interestingly this international study suggests that learners' performance was not much affected by teacher training but rather teacher quality. A Principal can very often identify a good teacher and a bad teacher but this is not necessarily related to qualification, making allowances however for a minimum level of subject knowledge. As regards the exam standards and whether they remain the same, he felt this area could be debated but this was in fact the particular task of Umalusi. He felt that they have done relatively well in maintaining standards although the methods used have not necessarily always been optimal. He added that this issue was not a major deficiency. Dr Bolton added that every year Umalusi takes a number of measures that relate in a complex way and triangulate to ensure that there is a consistency of results over time.

Another person from the audience made a further comment about the role of language in learning suggesting that there is very often a discrepancy between a learner's interpersonal communication skills and their cognitive academic language proficiency. She added that the assignments that learners are presented with in class do not adequately prepare them to access this language. The facilitator is often ill equipped to handle this task as they themselves do not know how to access this academic language proficiency; and this spirals down to learners. She felt it important to start equipping teachers with knowledge and skills needed to assess, beginning at the level of Blooms Taxonomy. She made a further point about data collection in schools, suggesting that it was very lacking. She suggested that data collection needs to be taken seriously by leadership in schools. That is, learners need to be measured more consistently and this information needs to be recorded and effectively analysed across grades if teachers are to become more accountable.

A representative of CEPD asked if it were possible to have a feedback session after this seminar that included all the different stakeholders at district and provincial levels to further discuss and understand the complexities of assessment. He felt this should be part of the feedback mechanism earlier discussed and would help to develop insight into why current practises are not helping to improve the quality of assessing, teaching and learning. He also expressed concern about contexts within which there were poor Principals, given the important role of leadership. He felt that this forum should be more inclusive at all levels so that people feel part of the process and can contribute to finding solutions.

The earlier speaker from the Department of Education stressed that it is very important for Umalusi to look at trends. She reiterated that the strategies discussed by Dr Cohen were first introduced in 2008 to try to influence the problematic gap. It has not yet been established whether school based assessment has got better or not and it is therefore premature to make

decisions about whether school based assessment is worth keeping or not. She felt one of the most important points made by Professor van der Berg was about how different schools are in South Africa and felt it important to have accountability without completely circumscribing teachers. She added that over the last five years 600 000 learners have been added in Grades, 10, 11 and 12. This is an unprecedented number internationally and it is not possible to expect government to produce the teachers for this volume of learners, yet South Africans are not sympathetic towards 'biting the elephant' one year at a time.

Dr Bolton concluded by saying that the enhancement of the system over time is incremental and suggested that every bit of research or discussion helps education to move forward. Ms Eugenie Rabie, Chief Operating Officer presented the two speakers with tokens of appreciation for their support of the work of Umalusi. In drawing the seminar to a close, Biki Lepota, Researcher with Statistical Information and Research Unit at Umalusi, thanked the participants for attending. He emphasised that it is through this form of dialogue that we will be able to build a sufficient knowledge base to improve the quality of education in the country. Further he announced proposed dates for the remaining three joint seminars.



37 General Van Ryneveld Street, Persequor Technopark, Pretoria
Telephone: +27 12 349 1510 • Fax: +27 12 349 1511
E-mail: info@umalusi.org.za • Web: www.umalusi.org.za

ISBN 062044875-X



9 780620 448758