

Has curriculum reform in South Africa really changed assessment practices, and what promise does the revised National Curriculum Statement hold?

SALOSHNA VANDEYAR AND ROY KILLEN

SALOSHNA VANDEYAR is a lecturer in the Department of Teaching and Training Studies in the Faculty of Education at the University of Pretoria. She specialises in Diversity and Education, Assessment Strategies and Practices, as well as Teaching and Training Studies. These topics also represent her field of interest and the foci of her research. She is currently pursuing research in the field of assessment, which is funded by the National Research Foundation of South Africa. The title of her proposed research is: A critical analysis of assessment practices in multilingual classrooms in South Africa

ROY KILLEN is an Associate Professor in the Faculty of Education and Arts at the University of Newcastle, Australia. He is also an Extraordinary Professor in the Faculty of Education at the University of Pretoria. He has worked extensively in South Africa over the past six years presenting seminars and workshops on assessment and outcomes-based education.

Abstract

This article examines the extent to which outcomes-based education, Curriculum 2005 and the Revised National Curriculum Statement provide guidelines for assessment that are consistent with principles of high-quality assessment. It illustrates that important principles such as reliability, validity and fairness are embodied in these curriculum frameworks, but that the principles are not always made explicit. It is claimed that this shortcoming is one of the reasons that concerns about why, how and when to assess learners have been evident in much of the debate surrounding recent South African curriculum reform.

The paper argues that if teachers understand the fundamental principles of high-quality assessment, then they will have little difficulty in adapting their assessment practices to the broad guidelines provided by OBE or to the specific guidelines provided by Curriculum 2005, the revised National Curriculum Statement or any future curriculum framework. The paper uses examples from case studies to suggest that when teachers ignore sound assessment practices, assessment becomes a meaningless activity divorced from learning.

Introduction

When outcomes-based education (OBE) was introduced in South African schools, it required teachers to follow some new approaches to planning, teaching and assessment. This was stressful for many teachers who felt that they were ill-prepared for this so-called paradigm shift, and who found it difficult to navigate through the maze of new jargon that accompanied OBE and Curriculum 2005 (Jansen, 1999; Department of Education, 2000, 2). There were many calls for the changes to be postponed until teachers had received adequate training and until schools had been provided with the required resources for this new way of teaching (Potenza & Monyokolo, 1999). As might have been predicted, the responses of teachers to the mandated changes varied in classical ways: a few teachers embraced the changes enthusiastically, many

reluctantly accepted the changes, and most resisted. One of the greatest areas of concern for teachers and, therefore, an area of strong resistance to change centred on the issues of assessment and reporting of learning. This article uses evidence from case studies in schools to examine how some teachers have responded to the demands for change and, in the process, have ignored the principles of sound assessment practices.

The article commences by outlining some of the basic principles of sound assessment practice. The principles are then used as a framework for examining typical assessment practices in pre-OBE schools in South Africa. From here the changes to assessment practices embodied in the principles of OBE and reflected in Curriculum 2005 and in the new National Curriculum Statement are examined. The paper explores the hypothesis that teachers' struggles with the 'new' approaches to assessment are grounded more firmly in their limited understanding of the principles of assessment than in any specific requirements of assessment in OBE. The paper concludes by suggesting some tentative solutions to this problem.

Principles of 'high quality' assessment practices

Regardless of the educational setting, high-quality assessment practices should satisfy certain common principles that are typically referred to as reliability, validity, fairness, discrimination and meaningfulness (Airasian, 2001; Gronlund, 1998; Herman, Aschbacher & Winters, 1992). When these principles are understood they provide a clear framework for all the major decisions that teachers need to make on assessment. When they are misunderstood or ignored, the resulting assessment practices are likely to result in the generation of worthless data. When teachers pretend that these data provide worthwhile indications of student learning, they are misleading themselves and greatly disadvantaging the learners in their classes. It is, therefore, appropriate to briefly review several of the key principles of assessment before exploring how they should have been considered in recent South African curriculum reform and how they should influence teaching practices.

Reliability

Reliable assessment items or tasks are those that are substantially free of errors of measurement. Because measurement errors produce inconsistencies, it is common to think of reliable tests as being those that produce consistent results despite apparent changes in the assessment situation (e.g. the test being administered at different times or the learner's performance being judged by different markers). Although various statistical procedures can be used to gauge reliability (Gronlund, 1998), such procedures are unusable for most teachers, either because teachers lack the knowledge necessary to apply them for analysing typical classroom tests, or because of time constraints and the impracticality of repeating tests. However, this does not mean that teachers should ignore the need for reliability. Rather, they need to try to minimise the extent to which learners' performances in each assessment task are influenced by unwanted variability arising from the learners (perhaps because they were hungry, tired or under stress) or the assessment task (perhaps because it was worded in a confusing way). Teachers also need to minimise the extent to which their judgements on learners' understanding are influenced by undesirable factors such as interruptions to marking or preconceived ideas about the learners' capabilities. For most practical purposes in schools, an assessment task can be considered reliable when the task, the conditions under which it is administered, and the marking are designed to minimise errors of judgement concerning learners' performance. If two teachers can compare their judgements, this is a bonus that will further minimise these errors. The basic question for teachers to consider is: "Have I tried to minimise the possibility that I will make errors of judgement about learners' performance?"

Fairness

For a test to be reliable, it must first of all be fair; it should not require learners to do unreasonable things or to do them under unreasonable circumstances. For example, it would be unfair to ask learners questions in a language they did not understand or to expect learners to answer an extremely large number of questions in a short time. The first step in achieving fairness in testing is for teachers to ensure that all learners have had a reasonable opportunity to learn the things that are being tested. Furthermore, the assessment strategies must be designed to ensure equal opportunity for success regardless of the individual learner's age, gender, physical or other disability, culture, language, socio-economic status or geographic location. The basic question for teachers to consider is: "Does the assessment task give every learner a reasonable opportunity to demonstrate his/her understanding or skill?"

Validity

Quite commonly, validity is taken to mean, "a test measures what it is meant to measure" (Hill, 1981, 22). However, Messick (1989) points out that such a narrow definition is really just an indication of the content relevance and content representativeness of the test – that is, a measure of whether each item in the test is relevant and a measure of whether or not the test as a whole samples an appropriate range of the content that learners have been expected to understand. Messick argues that validity should really be considered as an evaluative judgement on the degree to which there is evidence to support the appropriateness of the inferences that are drawn as a result of assessment. From this perspective teachers should not only be trying to maximise the validity of the tests they use; they should also be trying to maximise the validity of the inferences they make as a result of using those tests.

Teachers should consider whether their tests are assessing appropriate content (or outcomes), but they should also consider the special characteristics of the learners, the circumstances under which the test was administered and, most importantly, the theoretical and empirical evidence they have for reaching any conclusions on student learning. The basic question for teachers to consider is: "Based on the evidence provided by the assessment task, can I justify the conclusions I have reached about the achievements of each learner?"

Discrimination

Historically it has been considered important for tests and individual items to be able to distinguish or discriminate between learners who have learned whatever is being tested and those who have not. Typically, objective test items that did not distinguish adequately between respondents who scored high and low in the overall test were said to have a low discrimination index, and this was regarded as inappropriate. Unless teachers are using objective test items and unless they have the mathematical skills to perform the calculations, they will not be able to use this approach to discrimination. However, in a less structured way, it is still worthwhile for teachers to focus on the question: "Why does this test item elicit different responses from different learners, and are those responses indicative of the level of understanding of each learner?"

Meaningfulness and contribution to learning

Learners cannot be expected to make a serious attempt at an assessment task unless it is meaningful to them. If it is meaningful, the task will also have the potential to contribute to the students' learning. One of the rationales for more contextualised assessments is that they ensure

that students engage in meaningful problems that result in worthwhile educational experiences and higher levels of motivation (Herman *et al.*, 1992). To ensure that assessment tasks are meaningful, teachers must explain the purpose of assessment to learners, learners must see the tasks as realistic and worthwhile, and the teacher must deliberately link the assessment to important learning outcomes. Naturally, assessment tasks will not be meaningful to learners who do not have sufficient background knowledge or appropriate language skills. The key question for teachers in relation to this principle is: "Is the purpose of the assessment task clear to learners and will they understand how it will contribute to their learning?"

Assessment in pre-OBE schools

In order to investigate why many South African teachers are struggling to deal with the assessment requirements of outcomes-based education, it is necessary to examine how these requirements differ from the expectations embodied in the broad principles of high-quality assessment outlined above. A useful starting point is to consider the expectations regarding assessment that were imposed on teachers in the previous education system in South Africa.

Prior to the introduction of OBE, most South African schools adopted an approach to learning and assessment that placed a strong emphasis on the accumulation of isolated facts and skills. Assessment was generally separated from instruction and largely took the form of assessing discrete, isolated or fragmented knowledge and skills. Assessment in this paradigm was characterised by paper-and-pencil tests that emphasised academic exercises and the recall of textbook-based knowledge. The assessment criteria were rarely made explicit before learners attempted the assessment tasks. The assessment tasks were usually single occasion and single attribute in nature. Learners were assessed individually with much secrecy surrounding the tests (Department of Education, 1997b, 23). Assessment in pre-OBE schools was largely driven by the need to produce marks that could be recorded and reported to prove to the relevant authorities that assessment had taken place, rather than being an integral part of the learning process (Cockburn, 1997, 5). Consequently teachers generally did not consider assessment until after teaching had occurred. The shortfall of this approach is clear if we consider assessment to have a dual purpose – informing the teacher of learners' progress and providing a basis for reflection on teaching. Many teachers within this system simply assessed what they believed they had taught well, so that the learners' marks would reflect highly on their teaching ability. The focus of assessment was thus not on the growth and development of the learner, or in the interests of the learner.

Within this system assessment was largely summative, norm-referenced and judgemental in nature (Department of Education, 1997a, 3). Assessment had a short-term focus and it had very little to do with unlocking the potential of the learner. Assessment tasks emphasised content and factual recall, and "often entailed learning in parrot-fashion" (Cockburn, 1997, 5). The 'best learner' was the one who could most accurately reproduce the teacher's marking memorandum. Very little emphasis was placed on critical thinking skills or on valuing the personal inputs of the learner.

The format of assessment procedures was dictated by rigid bureaucratic structures that stipulated when and how assessment should be conducted. The major percentage of marks awarded to the learner was always based on written examinations (Cockburn, 1997, 5). Learners wrote what was called the June examination that constituted one-third of their final mark for the year. A final examination was written at the end of the year (November). In this system virtually no emphasis was placed on performance-based 'authentic' assessment.

Because this assessment system was largely norm-referenced, with a strong emphasis on competition among learners, learners could be ranked and these rankings could be reported to parents. This led to a common belief that performance relative to the class average was more important than learning in some absolute sense. Because norm-referenced testing was designed to highlight achievement differences among students, it also led to 'labelling' of students as being high or low achievers (Stiggins, 1994).

Promotion and retention were key elements of the pre-OBE school system in South Africa. Because promotion decisions were based primarily on end-of-year examinations, learners were placed under considerable pressure and stress in this system. There is little evidence to suggest that much consideration was given to the reliability and fairness of the examinations, and to the validity of the decisions made as a result of them.

In the pre-OBE system, reliability of assessment (apart from the matriculation examinations) was rarely considered because much of the testing was done 'once off'. Validity was not considered beyond content relevance and representativeness. Fairness was lacking because most of the examinations were 'once off' and, for many students, were in a language other than their home language. Under those circumstances it is doubtful that many students' results reflected their true abilities. The principle of discrimination was not necessarily designed into the tests, but was considered 'after the event'. The only form of meaningfulness and contribution to learning that was attached to most assessment tasks was that of production of marks and promotion to the next grade. The overall emphasis on the ultimate goal of obtaining the matriculation certificate at the end of twelve years of schooling distorted the purpose and focus of many assessment practices in schools. The shortcomings of the assessment practices in pre-OBE schools should have been addressed directly in the curriculum reforms driven by OBE, but this does not appear to have happened.

Assessment in OBE

When the South African school curriculum was reformed, the new approach was claimed to be outcomes-based (Department of Education, 1997a, 29). This so-called 'paradigm shift' brought with it many suggestions for changes to assessment practices. Before examining how these changes were built into Curriculum 2005, it is important to review the basic principles of OBE and their implications for assessment.

There is a strong argument that high-quality assessment practices in OBE are fundamentally no different from high-quality assessment practices in any other approach to education (Killen, 2002). To explore the soundness of this claim it is appropriate to consider the implications for assessment of each of Spady's (1994) four defining principles of OBE namely clarity of focus, designing down, high expectations and expanded opportunity for learning. The principle of **clarity of focus** requires that all assessment tasks must be clearly and explicitly linked to well-defined outcomes. These links are essential if the assessment is to produce evidence from which valid inferences can be made about learners' achievements. It can also be argued that the basic tenets of fairness require that learners are not assessed on things that they have not been helped to learn.

The **designing back** principle of OBE is based on the idea that each component of learning is included in a curriculum because it has been identified as contributing directly to learners' achievement of short-term outcomes that contribute to more complex outcomes that eventually lead to the exit outcomes of the programme. When this principle is applied to assessment, it requires that teachers need to be able to describe the purpose of each assessment task in terms of (a) how it provides information about learners' current understanding, (b) how it provides

information on learners' readiness to proceed to the next step in learning and (c) how it provides information on each learner's progress towards long-term outcomes. There is a clear link here to the more traditional notions of content validity and predictive validity (Messick, 1989) and to the commonsense notion that each assessment task should inform the teacher about the readiness of learners to proceed to more complex learning.

The OBE principle of **high expectations** is based on the idea that, given appropriate opportunities, all learners can achieve high standards. It reflects the idea that teachers' expectations, as well as their teaching practices, influence learners' achievements (Luke, Lingard, Ladwig *et al.*, 1998). Following this principle, assessment tasks must be challenging, not simply routine; the assessment must provide scope for learners to demonstrate deep levels of understanding and high levels of achievement; it must be possible to discriminate between low and high levels of achievement; and, excellence in student achievement must be recognised and rewarded. It is essentially because of this principle that OBE emphasises the use of criterion-referenced assessment rather than norm-referenced assessment. There are again links with the general principle of fairness in testing; if assessment is to be criterion-referenced then the criteria must be made explicit before the learners attempt the assessment task. This will also minimise errors and increase the reliability of the assessment.

The OBE principle of **expanded opportunity** embodies the idea that all learners can succeed if they are given adequate opportunity and time. What really matters is that learners are ultimately successful in their learning, not that they learn in a particular way or within a fixed period of time. Learners who do not achieve appropriately high levels of understanding at their first attempt must be provided with further opportunities to learn and to demonstrate their learning. Of course teachers have to work within practical constraints (e.g., learners attend school for a limited number of days each year), but they must also try to adapt to the needs of their learners. To implement this principle, teachers have to investigate alternative methods of assessment and to question their traditional approaches to issues such as assignment due dates. This OBE principle links most closely with the basic assessment principle of fairness. It is not fair to expect that all learners will learn (and be ready for assessment) in the same time. Nor is it fair to judge learners' achievements on the basis of a very limited number of opportunities to demonstrate what they have learned.

The relationship between the principles of assessment outlined earlier in this paper and the foundational principles of OBE can be summarised as follows (based on Killen, 2002):

1. The assessment procedures should focus clearly on the outcomes to be tested so that valid inferences can be drawn about learning.
2. The assessment procedures should be reliable. There should be a conscious effort to minimise measurement errors and allow learners to demonstrate their understanding at appropriate times and in ways that will produce consistent results.
3. Assessment procedures should be fair. The criteria for high-quality performance should be made explicit and the learners' opportunity to demonstrate their understanding should not be influenced by any irrelevant factors such as the learner's cultural background.
4. Assessment should reflect the knowledge and skills that are most important for learners to learn (that is, the building blocks for the achievement of long-term outcomes).
5. Assessment should challenge learners to the limits of their understanding and their ability to apply their knowledge. It will, therefore, discriminate between those who have achieved high standards and those who have not.

6. Assessment tasks should be authentic and meaningful so that they support every learner's opportunity to learn and, because learners are individuals, assessment should allow this individuality to be demonstrated.

From this brief overview it appears that there is nothing in the principles of OBE that contradicts the principles of high-quality assessment outlined earlier in this paper. It seems reasonable to assert that teachers who followed the principles of high-quality assessment prior to the introduction of OBE in South Africa would be required to make some procedural changes (such as using criterion referencing rather than norm referencing) but few substantial changes to their assessment practices after the introduction of OBE. To examine this assertion in more detail, it is necessary to consider the systemic framework that was established to guide the South African curriculum reforms.

Assessment guidelines in Curriculum 2005

Curriculum 2005 was an attempt to provide an outcomes-based framework for school education in South Africa. It engendered widespread efforts to transform the ways in which learners' work and learning were to be assessed. The most obvious change in assessment was in its general focus – away from a fixed body of content that was to be remembered towards a set of outcomes that were to be demonstrated. In this paradigm there was a strong emphasis on the application and use of knowledge in real-life contexts. Assessment was meant to be integrated with teaching and learning and to be cross-disciplinary in nature. Multiple opportunities to learn and to demonstrate learning were now acceptable, so opportunities for re-assessment became important.

Curriculum 2005 placed a strong emphasis on continuous assessment that was usually defined as "the assessment of the whole learner on an ongoing basis over a period of time where cumulative judgements of the learners' abilities in specific areas are made in order to facilitate further positive learning" (Le Grange & Reddy, 1998, 11). In the continuous assessment model, the principle of criterion-referenced assessment underpins all assessment; it is no longer necessary to compare a learner's performance with that of others. This was clearly a change from the pre-OBE focus on norm-referenced summative assessment and once-off, high-stakes examinations as a basis for promotion decisions. The ongoing formative assessment recommended in Curriculum 2005 was for the purpose of monitoring learners' progress, identifying their strengths and weaknesses and providing them with constructive feedback. These policy initiatives aimed to move the focus away from memorisation of content as an end in itself, toward a more thematic approach that encouraged learners to work with content in pursuit of greater understanding. Content was still important, but it was more important to translate content into meaning and meaningful action.

The assessment practices recommended in Curriculum 2005 certainly represented significant changes from the practices that were common in pre-OBE schools in South Africa. However, they were reasonably consistent with the 'ideal' OBE assessment practices outlined above. From this perspective, they did not represent departures from the basic principles of high-quality assessment practices. Reliability, validity, fairness, discrimination and meaningfulness were still the underpinning principles, even if they were camouflaged in a jungle of new jargon. This fundamentally important point was overlooked in much of the information that was provided to teachers. Documents from the Departments of Education (e.g. Cockburn, 1997) emphasised the paradigm shift and the new terminology (e.g. range statements) and almost totally ignored the fundamental principles such as reliability and validity. The focus on procedures, new terminology and the political rhetoric of transformation drew attention away from the need for teachers to understand principles such as validity. It is a sad fact that teachers cannot apply

assessment principles that they do not understand, as will be illustrated in the following case studies.

Case studies

The data reported here come from a pilot study of the assessment practices of Grade Four teachers in multilingual classrooms in South Africa. The study used just three teachers and was designed primarily to refine the observation, videotaping and interview techniques that are to be used in a larger study. All reported data were gathered by the principal author. Data from these teachers will be used to illustrate the thesis of this paper, namely that classroom assessment practices should be driven by the principles of high-quality assessment. Interviews with the three teachers established that they each believed that their assessment practices were aligned to outcomes-based education in general and to C2005 in particular. However, when the assessment practices of these teachers are compared with the principles of high-quality assessment practices, some major shortcomings can be observed.

The observations occurred over a two-week period. The three schools in this study were 'multicultural' in character, largely due to the desegregation of schools in South Africa. The first school (in which the teacher will be referred to as Teacher A) was an English-medium school catering exclusively for White children in the previous dispensation. The class observed contained 28 learners of whom 20 were White, seven were African and one was Indian. The second school, in which Teacher B taught, was an Afrikaans-medium school catering exclusively for White learners in the previous dispensation. The class observed contained 39 learners of whom 23 were African, 13 were Indian, two were White and one was Coloured. The third school, in which Teacher C taught, is an English-medium school that catered exclusively for Indian learners in the previous dispensation. The class observed contained 35 learners of whom 26 were African and 9 were Indian. English was the language of instruction used in all the observed classes.

The following descriptions summarise the activities observed during the lessons in which the three teachers were attempting to assess the ability of their learners to solve mathematics word problems. Teacher A gave her learners five mathematics word problems to solve within a time span of 40 minutes. She had spent two weeks prior to the assessment task building her learners up for the event by teaching them how to solve word problems, moving from the simple to the more complex. The assessment task was conducted as a follow-up activity to their investigation of the mathematics concepts of addition, subtraction and multiplication, which were embedded in the word problems. Because it was presented as a logical extension of their learning, and because the learners completed the task in a relaxed, friendly and comfortable environment, the assessment task appeared to create little stress for the learners.

Two of the questions that Teacher A asked were the following:

Grandpa collected 51 eggs. Unfortunately he dropped the basket and 27 broke. How many eggs did not break?

During the Boland Bank cycle race, the cyclists ride 187 km on the first day, 238 km on the second day and 91 km on the third day. How far do they ride in the three days?

During the last five minutes of the assessment task, Teacher A issued the learners with the following assessment criteria against which the learners' performance was to be assessed.

Table 1: Solving problems using a system

2 = Mastered;		1 = Acceptable	0 = Not acceptable	
1)	Writes an equation	2	1	0
2)	Puts known quantities into an equation	2	1	0
3)	Follows correct order of operations	2	1	0
4)	Shows all work for each step	2	1	0
5)	Displays work neatly, clearly and with the correct mechanics	2	1	0

The learners who completed the test early were allowed to take other activities from the activity rack at the back of the classroom and silently to continue working on those activities.

When Teacher A's assessment practices are compared with the principles of high-quality assessment outlined earlier in this paper, several important points can be noted. Each question was relevant to the outcomes that Teacher A was trying to assess, and the total 'coverage' of the questions seemed appropriate since they dealt with all the operations of addition, subtraction and multiplication. To this extent, she had established a basis for drawing valid conclusions concerning the learners' ability to solve these types of problems. The questions were at an appropriate level of difficulty and the test conditions were fair. However, there were several shortcomings in this assessment exercise.

The manner in which the questions were phrased indicates some bias in the assessment task and this compromised both the reliability and the validity of conclusions drawn from the results. The questions were phrased from a particular cultural background (e.g., both the words 'Grandpa' and 'Boland Bank' are affiliated with the Afrikaans culture) and wording such as this could alienate learners from other cultural backgrounds. The problems might have had more meaning for the learners had they described situations within the life experiences of the learners.

Because the marking criteria were not revealed until most learners had completed the test, some learners may not have focused on following the five steps that would lead to full marks for each question. Because the marking guide placed a heavy emphasis on procedures rather than understanding, learners were penalised heavily for making 'mechanical' errors even if they had a deep understanding of the concepts that the problems were testing. This compromises the validity of the inferences that the teacher was drawing from the learners' results.

Teacher A's approach to assessment could be summarised as follows: The problems she used were relevant to the outcomes she was trying to test and she attempted to integrate assessment with learning. However she seems to have displayed a traditional concern for procedural correctness rather than for understanding. There was no obvious concern about fairness and reliability, or about the validity of her interpretation of the assessment results.

Teacher B conducted her assessment task under rigid conditions. The learners were expected to attempt the 'test' in total silence and under strict invigilation from the teacher. The test was scheduled for 40 minutes and learners were instructed that if they completed the test sooner, they had to put their heads on the table and sleep. Observations of the teaching of mathematics word problems in this class in the week prior to the test suggested that learners needed more work in this section before formal assessment occurred. However Teacher B had decided that it was time for assessment. The assessment task consisted of five questions, of which three were not word problems. One involved completing number patterns; another involved converting

Roman numerals to Arabic numerals and the third dealt with classification of objects as two-dimensional or three-dimensional. The two word problems were more complex than those used by Teacher A and involved interpretation of tabular data. One of these problems is shown below.

Teacher B's justification for the first three questions (that were not word problems) was that she was trying to compensate for the language barrier. She was aware that some of her learners had a problem with English and that they would not be able to perform well in word problems. She did not appear to consider this with the two (relatively difficult) word problems.

There are several obvious shortcomings in the approach to assessment taken by Teacher B. Because most of the questions were not testing the learners' ability to solve word problems (or


Figure 2: Exemplifying shortcomings in a teacher's approach to assessment

Question 4

4. At Atlantic Primary School there are 300 children in Std 3, 200 in Std 2 and 400 in Std 1.

This table shows the kind of transport that the children use to get to school.

	bus	train	bicycle
Std 3	5	200	40



Answer these questions:

- How many children in Std 3 use transport? _____
- How many children in Std 2 use transport? _____
- How many children in Std 1 use transport? _____

The rest of the children in these classes walk to school.

- How many children in Std 3 walk? _____ (6) _____
- How many children in Std 2 walk? _____
- How many children in Std 1 walk? _____

Source: Material anonymously supplied by Teacher B

the underlying ability to perform addition, subtraction or multiplication) and since some of the learners had extremely poor language skills, the validity of any inferences drawn from the learners' results has to be questioned. The assessment tasks were certainly not fair to the learners with poor language skills and it is quite likely that they were meaningless.

Teacher B's approach to assessment could be summarised as follows: She seemed to have no concern for relating questions to concepts that were supposed to be tested, and her concern about the 'language barrier' was just a superficial response to the language difficulties of her students. She seemed unaware that she was disregarding the principles of fairness and reliability, and that this made her conclusions about student learning invalid.

The assessment practices of Teacher C provide some indication of why teachers are struggling to implement outcomes-based assessment. Teacher C had not prepared her learners for the assessment event; her assessment task was driven by the need to produce marks for the end-of-term report card. (It was the week before the end of term and no other assessment tasks had been given to the learners.) The assessment task was conducted under chaotic circumstances. Many of the learners did not know what was expected of them and the environment had been set up for widespread copying to take place. A few of the learners did not speak English and others had a very limited proficiency in English.

Teacher C began by dividing the learners into groups ("because that's what should happen in OBE") and nominated an Indian student as the leader of each group ("because they could read

English"). She then issued the following group assessment criteria to each group and read every question to the class.

Table 2: Group assessment – word problems

Name of group leader: _____	
Other members in the group:	
_____	_____
_____	_____
_____	_____
1.	Were all the questions answered? _____
2.	Which question was the easiest? _____
3.	Which question was the most difficult? _____
4.	Did the group look for any clues? _____
5.	Were the clue words and numbers underlined? _____
6.	Did the group read each problem? _____
7.	Who was very helpful in the group? _____
8.	Who made the most noise in the group? _____
9.	Did every person help with the problem in the group? _____
10.	Did it take very long to solve the problem? _____
11.	Did we waste a lot of time? _____
12.	What should we do next time to work better in a group? _____

She subsequently gave each of the learners a worksheet based on mathematics word problems similar to those used by Teacher A and instructed them to work in their groups to solve the problem. Learners were confused as to whether they were being assessed on their participation in the group or on the mathematics problems. Within each group one or two learners were writing down answers on the worksheet. The others copied, sometimes incorrectly. Some learners appeared to be totally lost. The teacher then collected the worksheets and assigned individual marks to each of the learners without considering the group assessment criteria list at all. Examples of the maths word problems used were:

The tuckshop sells 168 crisps and 249 cooldrinks. How many more cooldrinks than crisps were sold?
 Sam is 9 years older than his sister Lisa. If Lisa is 16 years old, how old is Sam?

The assessment practices of Teacher C raise some serious concerns. Because most learners simply copied the work of those who appeared to know what they were doing, their results could not be considered a reliable indication of their understanding of the mathematics concepts incorporated into the word problems. This problem was compounded because most learners were under the impression that they were being assessed on group dynamics, not their mathematics knowledge. The task was simply not fair to the learners who had limited English proficiency, and probably had no meaning for them. As a consequence, no valid inferences could be drawn about the learners' understanding.

Teacher C's approach to assessment could be summarised as follows: Her primary concern was to generate marks. She seemed unaware that the process she was using made those marks virtually meaningless. She was under the misapprehension that group work is essential in OBE. The confusion she caused by overemphasising the group work marking criteria compromised the fairness and reliability of her testing. There was no justification for claiming that her interpretations of the test results were valid.

The three teachers in this case study seemed unaware of (or else they deliberately ignored) some of the basic principles of assessment as espoused by writers such as Airasion (2001). If learners do not know what is expected of them in an assessment task, or they do not understand the criteria by which their performance will be judged, they cannot try to satisfy those criteria. It is impossible to get reliable (consistent, error-free) data from learners who cannot understand the questions they are asked. Questions that are incomprehensible (for reasons of language) or difficult to understand for cultural reasons, lack the meaning that is necessary for serious engagement. In these circumstances the evidence obtained from the assessment task simply cannot provide a sufficient basis for drawing valid conclusions about learners' understanding.

We are not claiming that the actions of the three teachers described above are representative of the approach to assessment taken by other teachers in South Africa. However, these three brief examples illustrate that when teachers ignore (or do not understand) the basic principles that define high-quality assessment practices, the value of the results that they obtain from learners is seriously compromised. There is a grave danger that in these circumstances the teachers' lack of awareness of the illegitimacy of their practices can lull them into a false belief that what they are doing is worthwhile and in the best interests of their learners. When this happens, assessment ceases to be an integral part of learning and becomes a pointless exercise of generating meaningless marks to support the pretence that learning is being facilitated. This problem is amplified when teachers hide behind a lack of understanding of jargon, such as the terminology used to describe assessment in Curriculum 2005, and do not address their more important lack of understanding of the principles of sound assessment. The Revised National Curriculum Statement provides an opportunity for this crucial problem to be addressed.

Assessment guidelines in the National Curriculum Statement

A Ministerial Committee reviewed Curriculum 2005 and its implementation in 2000. The Review Committee recommended that strengthening the curriculum required streamlining its design features and simplifying its language through the production of an amended National Curriculum Statement. It further recommended that this Revised National Curriculum Statement should reduce the curriculum design features from eight to three: Critical and developmental outcomes, learning outcomes and assessment standards (Department of Education, 2000, 4). The revised National Curriculum Statement is not a new curriculum but a streamlining and strengthening of Curriculum 2005. It keeps intact the principles, purposes and thrust of Curriculum 2005 and affirms the commitment to outcomes-based education. Introduction of the revised National Curriculum Statement Grades R-9 (Schools) in the Foundation Phase is planned for 2004.

The Revised National Curriculum Statement Grades R-9 (Schools) released in May 2002 (hereafter referred to as the NCS) introduced several important changes to proposed assessment practices in schools. The most significant change was a shift from the criterion-referenced assessment that was recommended in C2005 to a form of standards-referenced assessment. This system is intended to operate as follows: In each Learning Area in each Phase of schooling there is a set of outcomes that define what learners are expected to achieve – these *phase outcomes* provide a direct link to the OBE principle of clarity of focus. For each grade of schooling, there is a set of *assessment standards* that define the levels of knowledge, skills and attitudes that learners will be required to demonstrate as evidence that they have achieved each phase outcome to an appropriate depth and breadth. This means that in each Phase the outcomes remain the same from grade to grade while assessment standards change from grade to grade. Because these standards are grade-specific, they describe how conceptual understanding is meant to progress in each Learning Area.

This change to describing intended learning in each Phase in terms of progressing standards rather than changing outcomes has important implications for teaching and assessment. For the first time in South Africa, teachers are being encouraged to think of outcome attainment as a continuum of possibilities, rather than as a dichotomy. They are no longer being asked to put learners into categories of 'achieved/not achieved' for each outcome. Instead they have to think about how well each learner has achieved each outcome. This is the approach to assessment that had been advocated by Killen (2000). Unfortunately, the National Codes for reporting learner achievement introduced in the NCS still require learners' performance of each outcome for each Grade to be recorded on a four-point scale (*exceeded, satisfied, partially satisfied or not satisfied*). This reluctance to move to a fully 'standards-referenced' approach to assessment is also reflected in the requirement that reporting should also be normative and "... contain comments on the learners' performance in relation to peers" (Department of Education, 2000, 100).

The NCS places a new focus on assessment when it states that "the main purpose of assessing learners is to enhance individual growth and development, to monitor the progress of learners and to facilitate their learning" (Department of Education, 2000, 94). (Note: The section on assessment is generic to all the learning area documents of the Revised NCS. For convenience we have quoted from the Mathematics document.) This emphasis on the formative role of assessment and on its integration with teaching and learning is a clear departure from the pre-OBE school system in which assessment had the prime function of determining learners' readiness to progress. It strengthens the Curriculum 2005 idea that assessment can actually help students to learn; thus it should be seen as a 'refinement' rather than a 'new direction'.

The specific assessment guidelines provided for teachers at the end of each Learning Area Statement explicitly attempt to link assessment practices to the principles of outcomes-based education. The guidelines open with the statement: "The assessment framework of the Revised National Curriculum Statement Grades R-9 (Schools) is based on the principles of outcomes-based education" (Department of Education, 2000, 93). When defining the characteristics of continuous assessment, the NCS claims that this approach to assessment "covers all the outcomes-based education principles" (Department of Education, 2000, 95). Although these principles have not been enunciated in the NCS, it is reasonable to assume that they are the principles outlined by Spady (1994), that is the principles of clarity of focus, designing down, high expectations and expanded opportunity.

The claim that the guidelines follow these principles is certainly justified in relation to the principle of *clarity of focus* because assessment is very clearly linked to outcomes, and the assessment standards clarify what is expected in the demonstration of each outcome. There is also the specific direction that "the methods chosen for assessment activities must be appropriate to the Assessment Standards to be assessed and the purpose of the assessment must be clearly understood by all the learners and teachers involved" (Department of Education, 2000, 95). In a less direct way the importance of clarity of focus is reinforced through the suggestion that there should be "agreement between teachers in the same grade about what is considered necessary to satisfy the Learning Outcomes" (Department of Education, 2000, 96).

In relation to the principle of *expanded opportunity*, the guidelines are also reasonably explicit. They suggest that "All learners do not need to be assessed at the same time and in the same way" (Department of Education, 2000, 95), thus allowing for the possibility of multiple and varied opportunities for learners to demonstrate their achievements. This idea is reinforced by emphasising that choices involving assessment strategies are subjective and "dependent on the teacher's professional judgement" (Department of Education, 2000, 95).

In relation to the principle of *designing down*, the guidelines are less clear. Certainly, the assessment standards in each phase show a progression from relatively low levels of achievement of each outcome to higher levels of achievement. However, it is not clear whether these standards were designed down or designed up. There is also a claim that the NCS "contextualises the Critical and Developmental Outcomes within the Learning Outcomes and Assessment Standards" (Department of Education, 2000, 94), but the significance of this for teachers' assessment practices has not been explained.

In relation to the principle of *high expectations*, the guidelines are less explicit. There is a general reference to helping learners to "reach their full potential" (Department of Education, 2000, 94), and to the idea that continuous assessment "supports the growth and development of learners" (Department of Education, 2000, 95), but there is no specific reference to the idea that, given appropriate learning opportunities, all learners can achieve high standards.

The NCS places a strong emphasis on the administrative aspects of assessment. Whereas C2005 had assumed that administrative aspects are in place in all institutions, the NCS prescribes minimum requirements in assessment that are necessary for all schools to function. For example the NCS specifies the essential details that must be included on report forms. While this information will be useful, particularly the requirement that each learner's "strengths and needs" should be reported, there are no guidelines that would enhance the validity or reliability of the information that will be reported.

Another important change in the NCS is the suggestion that assessment of learner performance should be a routine part of "monitoring the performance of the education system" (Department of Education, 2000, 94). The NCS recommends that "systemic assessment" be undertaken at the end of each phase, i.e. Grade 3, 6 and 9 within the General Education and Training Band. This "systemic assessment" would inform the education system of the level of performance of the learners and thus the level of performance of the education system. The NCS suggests that this "systemic assessment" be conducted on a representative sample of schools.

The NCS retains the C2005 idea of recording learner "case histories" in learner profiles. The NCS suggests that educators should use these profiles as an aid/tool in an attempt to understand their learners as they progress from one year to the next. The learner profiles represent the cumulative records of a learner, ranging from his/her personal details to emotional and social behaviour, report cards, medical certificates, assessment records, etc.

This brings us to a consideration of whether or not the principles of assessment outlined in the NCS have a solid foundation in the basic principles of assessment outlined in the first section of this paper. It can certainly be argued that the NCS incorporates most of the principles of assessment embedded in the principles of OBE. To the extent that these principles account for the fundamental issues of reliability, validity and fairness, we can imply that these principles are embedded in the NCS. However, the NCS makes no direct reference to the foundations of sound assessment practices, other than the statement that assessment should be "valid, reliable, fair, learner-paced, and flexible enough to allow for expanded opportunities" (Department of Education, 2000, 94). Validity, reliability and fairness are not defined in the Curriculum and Assessment Glossary attached to the NCS, implying that teachers will know what these terms mean and how they can be used to guide assessment practices. This is a major weakness of the NCS guidelines on assessment because past experience has shown that teachers do not necessarily understand how to build these characteristics into their assessment practices.

In summary the NCS makes several major advances in helping to simplify assessment in South African schools. It removes much of the confusing jargon of C2005, it makes explicit links to the principles of OBE and it provides a workable standards framework. Its major weakness is that teachers are still not provided with guidelines on the fundamental principles of good assessment practices.

Conclusion

The fundamental principles of assessment that are usually referred to as reliability, validity, fairness, discrimination and meaningfulness were established long before South Africa embarked on its post-1994 curriculum reform journey. This paper has attempted to show that OBE, Curriculum 2005 and the Revised National Curriculum Statement all address these basic principles of high-quality assessment, albeit in different ways. None of these curriculum frameworks introduce fundamentally new ideas about assessment. Instead, they each suggest slightly different ways of building the principles of high-quality assessment into curriculum and teaching practices. These suggestions do not always make explicit reference to principles such as validity but, as shown in this paper, the principles appear to be embedded in the guidelines.

When South Africa began its curriculum reform, the 'paradigm shift' that was so frequently advocated was, at least in relation to assessment, a misleading idea. There was no need for a totally new way of thinking about assessment. The basic principles of assessment did not change with the introduction of OBE, Curriculum 2005 or the Revised National Curriculum Statement; nor did the fundamental principles that govern effective teaching and learning. There were changes in emphasis (for example, from summative to formative assessment) but these should have been easily accommodated by teachers who understood assessment principles such as reliability and validity (principles that should have been emphasised in pre-OBE education in South Africa). The confusion and resentment that arose because of the misguided emphasis on a 'paradigm shift' were inevitable. A more productive approach to reform would have been to encourage educators at all levels in South Africa to adhere to a set of basic principles of teaching and assessment that would ensure that their assessment practices are reliable, fair and meaningful so that the inferences and decisions they make as a result of those assessment practices are justifiable and valid.

Teachers cannot be expected to apply assessment principles that they do not understand. Therefore, those who propose curriculum change have an obligation to ensure that the principles driving those reforms are explicit and that they are explained clearly. By removing much of the jargon that characterised Curriculum 2005, and by making the links to OBE more explicit, the NCS has taken two important steps in helping teachers to use more appropriate assessment practices. Unfortunately, the NCS falls short of providing information and explicit guidelines that would help teachers to focus specifically on the fundamental principles of high-quality assessment practices. This shortcoming reduces parts of the NCS to a 'recipe' for assessment – a set of administrative requirements that must be followed without understanding the principles on which those procedures are based. Such an approach in the NCS perpetuates the perception held by many teachers that assessment is a matter of technical procedure (something that must be done to satisfy the bureaucrats), rather than a matter of professional judgement (something that should be done to help students learn). This approach also perpetuates the view that teaching is a technology rather than a complex social process.

References

- Airasian, PW. 2001. *Classroom assessment: Concepts and applications*. 4th ed. Boston: McGraw Hill.
- Cockburn, P. 1997. *Building a brighter future. Curriculum 2005 – Questions and answers*. Pretoria: Gauteng Department of Education.
- Department of Education. 1997a. *Assessment*. Pretoria: Department of Education.
- Department of Education. 1997b. *Assessment policy in the general education and training phase. Grade R to Grade 9 & Adult Basic Education*. Pretoria: Department of Education.
- Department of Education. 1997c. *Curriculum framework for general and further education and training*. Pretoria: Department of Education.
- Department of Education. 2000. *Report of the Review Committee – Executive summary, 31 May*. Pretoria: Department of Education.
- Department of Education. 2002. *Revised National Curriculum Statement R-9 (Schools)*. Pretoria: Department of Education.
- Gronlund, N. 1998. *Assessment of student achievement*. 6th Ed. Boston: Allyn and Unwin.
- Herman, JL, Aschbacher, PR & Winters, L. 1992. *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hill, J. 1981. *Measurement and evaluation in the classroom*. Columbus, OH: Merrill.
- Jansen, JD. 1999. Why outcomes-based education will fail; An elaboration. In *Changing curriculum: Studies on outcomes-based education in South Africa*, Jansen, J & Christie, P. (eds). Kenwyn: Juta & Co.
- Killen, R. 2000. *Standards-referenced assessment: Linking outcomes, assessment and reporting*. Keynote address at the Annual Conference of the Association for the Study of Evaluation in Education in Southern Africa, Port Elizabeth, South Africa, 26-29 September.
- Killen, R. 2002. Outcomes-based education: Principles and possibilities. *Interpretations*, **35**(1), 1-18.
- Le Grange, L & Reddy, C. 1998. *Continuous assessment: An introduction and guidelines to implementation*. Kenwyn: Juta.
- Luke, A, Lingard, B, Ladwig, J, Mills, M, Hayes, D & Gore, J. 1998. *School reform longitudinal study: Interim report*. [Submitted to Education Queensland by the Graduate School of Education, The University of Queensland]. Brisbane, Australia: University of Queensland.
- Messick, S. 1989. Validity. In *Educational measurement*, 3rd ed., Linn, RL (ed.). New York: Macmillan.
- Potenza, E & Monyokolo, M. 1999. A destination without a map: Premature implementation of Curriculum 2005. In *Changing curriculum: Studies on outcomes-based education in South Africa*, Jansen, J & Christie, P. (eds). Kenwyn: Juta & Co.
- Spady, W. 1994. *Outcome-based education: Critical issues and answers*. Arlington, VA: American Association of School Administrators.
- Stiggins, RJ. 1994. *Student-centered classroom assessment*. New York: Merrill.