

Assessing the attainment of learner outcomes in Natural Science of the New South African Curriculum

A. Hattingh¹, J.M. Rogan², C. Aldous³, S. Howie¹ and E. Venter¹

¹University of Pretoria, South Africa; ²Hiroshima University, Japan;

³University of KwaZulu-Natal, South Africa

annemarie.hattingh@up.ac.za

Abstract

The second phase of the Mpumalanga Secondary Science Initiative (MSSI) was launched in Mpumalanga Province, South Africa in 2003. The MSSI seeks to improve the teaching and learning of mathematics and science in all secondary schools in the province over a period of three years. To achieve this goal an in-service system has been developed. The long-term research of the MSSI is aimed at examining the effect of the intervention on the three science learning outcomes as defined by the new curriculum. Tests were developed to assess 10 learners in each of Grades 8 and 9 in 40 schools, representing a 6% sample. This paper reports on the methodology of the baseline assessments undertaken and the extent of learners' attainment levels in the three Natural Science outcomes. It then looks at the results in terms of learners' socio-economic backgrounds. Finally it discusses some of the dilemmas encountered in this research, such as the issue of validity in the light of possible misalignments between the instrument, practice and policy.

Introduction

Curriculum 2005 (C2005) (Department of Education, 1997) unequivocally embraces outcomes-based education (OBE). The rationale is that for too long South African learners have memorised content, which they are required to regurgitate in tests and examinations. With the introduction of OBE, the focus shifts to what learners can do with their knowledge, and in particular whether they can use what they know to meet the specified outcomes. The C2005 (Department of Education, 1997) document makes the following claims:

- The move towards an outcomes-based approach is due to growing concern around the effectiveness of traditional methods of teaching and training, which were content-based. An outcomes-based approach to teaching and learning, however, differs quite drastically and presents a paradigm shift. According to Spady (1994) outcomes are high-quality, culminating demonstrations of significant learning in context.
- An outcomes-based education and training system requires a shift from focusing on teacher input (instructional offerings or syllabuses expressed in term of content) to focusing on the outcomes of the learning process.
- Outcomes-based learning focuses the achievement in terms of clearly defined outcomes, rather than teacher input in terms of syllabus content.
- In outcomes-based learning, a learner's progress is measured against agreed criteria. This implies that formal assessment will employ criterion-referencing and will be conducted in a transparent manner.

The new curriculum defined eight Learning Areas for Grades 1 to 9, one of which was the Natural Sciences. Nine specific outcomes were created, along with range statements and assessment criteria. Before long, it became apparent that the way in which C2005 had been framed was too complex and cumbersome. Committees were appointed to "revise and streamline" the original documents. In this revised version for the Natural Sciences, the original nine specific outcomes were reduced to the three shown below (Department of

Education, 2002). Grade level assessment standards were subsequently devised for each of these three outcomes.

Table 1: The three Natural Science outcomes of South Africa’s Curriculum 2005

<u>Learning Outcome 1 (LO1):</u> Scientific investigations	Learners act confidently on their curiosity about natural phenomena; they investigate relationships and solve problems in science, technology and environment contexts.
<u>Learning Outcome 2 (LO2):</u> Constructing knowledge	Learners know, interpret and apply scientific, technological and environmental knowledge.
<u>Learning Outcome 3 (LO3):</u> Science, society and environment	Learners are able to demonstrate an understanding of the inter-relationships between science and technology, society and the environment.

Purpose

The purpose of this paper is to report on the initial results of a project to assess achievements of learners in Grades 8 and 9 in the three outcomes in Table 1. All the Grades 8 and 9 learners in the sample population were exposed to C2005. While these results will be used as a baseline for an ongoing project, they are of interest in their own right as they represent one of the first attempts in South Africa to assess the new Natural Science outcomes on a large scale, and to make cross-grade comparisons. We will first present the results as raw data. Second, we will look at the results in terms of learners’ socio-economic backgrounds. Finally we will discuss some of the dilemmas encountered in this research, such as the issue of validity in terms of possible misalignments between the instrument, practice and policy.

Context of the study

The study described here is located within Mpumalanga, one of the nine provinces in South Africa. It is mostly a rural province, without large cities, situated on the eastern border of South Africa. The research is part of the Mpumalanga Secondary Science Initiative (MSSI), which has been active in the province since 1999. MSSI is a project of the Mpumalanga Department of Education, supported financially and professionally by the Japanese International Cooperation Agency (JICA) and by the Universities of Hiroshima, Naruto and Pretoria. During the first three-year phase it sought to develop a province-wide in-service system, with particular emphasis on school-based or cluster-based professional development. The professional development programme was designed by all stakeholders in the MSSI partnership. Its focus was on building capacity of those responsible for professional development of teachers.

The second three-year phase, launched in 2003, focuses on delivery at the classroom level. The capacity building of phase one is now expected to make a difference in the quality of teaching and learning of science and mathematics (e.g. see Howie, 2002; Shavelson, McDonnell & Oakes, 1987). In other words, the MSSI seeks to create support for a greater fidelity in the implementation of the intended curriculum in all secondary schools in the Province. A key part of this effort includes the development of learning support materials by participants who include teachers as well as Curriculum Implementers (CI’s) that can be used in the science and mathematics classes.

Theoretical underpinning of the study

The conceptual model used in this research is largely based on school improvement research, (e.g. see Fullan, 1999; MacBeath, 1999; MacBeath, Schratz, Jakobsen & Meuret, 2000). The

model is represented in Figure 1. In this paper, we consider the effect of the intervention on the performance of learners in Science Learning Outcomes 1 to 3 of the National Curriculum.

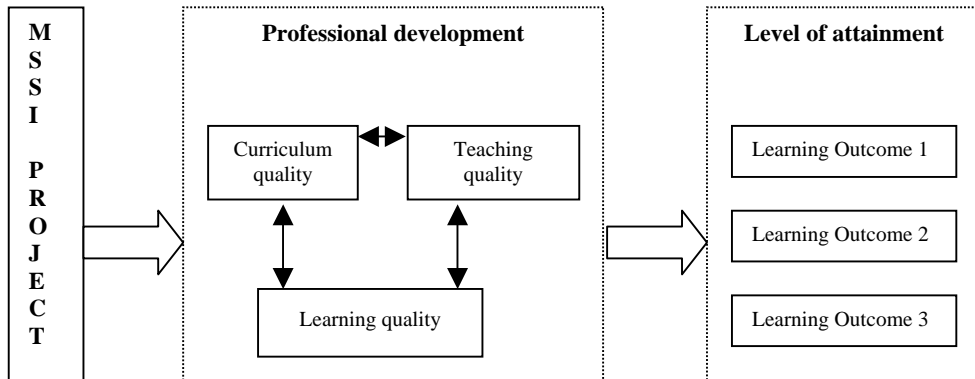


Figure 1: Conceptual framework for MSSI assessment

In one sense, the assessment of outcomes, especially those involving some kind of a performance, by means of a multiple-choice test, is a contradiction in terms. There is a vast difference between actually planning and doing a science experiment and selecting one of four alternatives on a test. Nevertheless, it is possible to frame multiple-choice questions that probe aspects of designing research questions and designs. Although less than desirable, such questions may well be the only way to provide some form of standardised, large-scale assessment, where such information is appropriate and required. Lawrenz, Huffman, & Welch (2001) tested 3 500 Grade 9 learners in the USA using four different formats: multiple-choice, open-ended, hands-on laboratory task and a full-scale investigation. The correlations between the multiple-choice and open-ended questions were found to be 0.95 and 0.96 respectively. (There were two alternative forms of these tests.) Correlations between the multiple-choice and the hands-on task were 0.59 and 0.68 respectively. The figures given here are for the overall results. When the same correlations were investigated for low-, middle- and high-scoring learners, it was found that they were higher for the high-scoring learners. In other words, the multiple-choice items are reasonably good predictors of both the open-ended questions and for the hands-on task, especially for the top learners. Whether these correlations would hold in South Africa is an open question, bearing in mind that by and large our learners are low-achievers by international standards. For the purposes of this paper, we make the assumption that the above findings can be generalised in South Africa.

Method

The instruments

Our research required an assessment procedure that was easy to conduct and analyse since a large sample was necessary. Guided by Lawrenz, Huffman, & Welch (2001), we opted for pencil and paper tests, especially using multiple-choice formats. While large-scale tests such as TIMSS (Third International Mathematics and Science Study) and PISA (Programme for International Student Assessment) have been developed and are available, they were not designed to fit South Africa, the National Curriculum Statement, or our research purposes. We chose to develop our own instruments.

The first step was to assemble a panel of science educators to generate a bank of items. These items were based on the outcomes in Table 1 and the assessment standards generated by grade

level from them. We decided that about 80% of the items should be multiple-choice. For the outcome LO1, investigating, we supplemented the multiple-choice questions with open-ended questions. Three to four times more items were prepared than could be used in the final version of the tests. All items were piloted at three high schools, and then subjected to item analysis.

The following criteria were used to select items for the final versions of the tests.

1. Items that were too difficult i.e. whose item difficulty fell below 25-30% were marked for elimination.
2. Items whose discrimination index fell below 0.2 were marked for elimination. Since the reliability of a test depends on high correlations between items and the overall total, a test with items having a higher discrimination index will have a higher reliability.
3. Points 1 and 2 above notwithstanding, the final tests needed to be balanced in terms of outcomes and content. Short answer open-ended items were retained despite the fact that by and large they were very poorly answered in the trials.

Based on the item analysis, the final version of the Grade 8 and 9 Natural Science test was developed, the composition of which is shown in Table 2.

Table 2: Number of items per outcome

Learning outcome	Natural science Grades 8 and 9 items
1	10 multiple-choice 1 five section open response
2	10 multiple-choice
3	10 multiple-choice

Validity and reliability

The central claim for the validity of the test rests on the correspondence of items to the three Natural Science outcomes of the revised C2005 and their associated assessment standards. One third of the multiple-choice items were devoted to each of the outcomes. The open-ended items were based on the first outcome only. Hence the validity claim is made in terms of the intended curriculum only, and not the implemented one. The primary purpose of the study was to collect base-line data on the extent to which C2005 is being implemented over a three-year period, rather than to assess the competencies of individual or groups of learners. The consequences of this choice are discussed later in the paper.

The reliability of the overall test was calculated using the Cronbach alpha and found to be 0.62. Although the test was designed to assess three different outcomes, the reliabilities of the sub-sections of the test were of the same order as that of the test as a whole. Hence it would appear that the test as a whole is as reliable in assessing a single construct (general science ability) as its parts are in assessing the three different outcomes. Given that subsequent analysis of the items suggested a high degree of guessing, this figure was higher than expected.

Data collection

There are about 658 secondary/high schools in Mpumalanga. A 6% (40 schools) sample was randomly selected. The grade appropriate tests were administered to between 20-24 randomly selected learners in each of Grades 8 to 12 in each of the 40 schools. In the case of Grades 8 and 9, half of the learners wrote the science test and the other half the mathematics test. For this paper, we analyse the results of the ten Grade 8 and 9 learners from the 40 randomly selected schools who wrote the Natural Science test. Grade 8 and 9 learners wrote the same test.

Data collection occurred during the second and third weeks of September 2003. In all cases, tests were administered by trained field-workers from the University of Pretoria who followed a standardised approach. Each team visited between two and five schools.

The test cannot be characterised as a high-stakes instrument. It is unlikely that any learners actually studied for the test. Indeed, many may have been unaware that they would be writing a test until the actual day in question. This situation does imply that the test was able to assess what the learners really knew, as opposed to what they might be able to memorise (and perhaps subsequently forget) by cramming techniques.

Results

The achievement levels on the test

The Natural Science mean score (converted to a percentage) for the 710 Grade 8 and 9 learners tested was 26.5% for the test as a whole, while it was 34.4% if only the multi-choice questions were considered. The frequency distribution of the total scores for the multiple-choice items is shown in Figure 2. The distribution is positively skewed, indicating an overall low performance. Total scores have been converted to percentages and grouped into ranges of 10%.

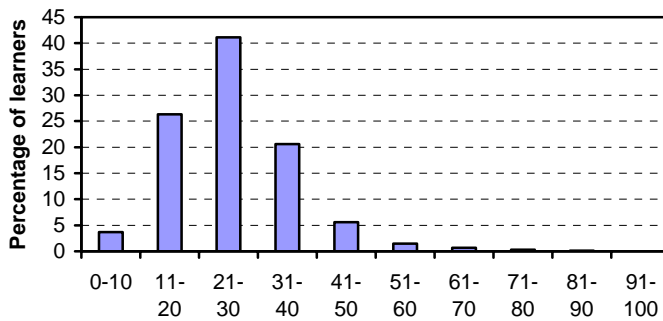


Figure 2: Frequency distribution of the Grade 8 and 9 Natural Science scores

The results for the multiple-choice section of the test, overall and by outcome, are shown in Table 3.

Table 3: Results for Grade 8 (n=352) and 9 (n=358) overall and by outcome

	Outcome 1			Outcome 2		
	Mean	t-test	p-value	Mean	t-test	p-value
Grade 8	29.6 %	3.584	0.00036 *	32.3 %	0.524	0.60079
Grade 9	34.2 %			32.9 %		
	Outcome 3			Total		
	Mean	t-test	p-value	Mean	t-test	p-value
Grade 8	35.9 %	4.122	0.00004 *	32.6 %	3.866	0.00012 *
Grade 9	42.0 %			36.3 %		

* p < 0.01

The overall low performance in all three learning outcomes by Grades 8 and 9 is evident from the data. For outcomes 1 and 3, as for the total, there were significant increases in performance from Grade 8 to Grade 9. There was no significant difference in LO2, science knowledge.

The open-ended questions were poorly answered. The results of some of the items (concerning LO1, science processes) are given below. The question presented a situation describing a

potential cause/effect relationship, the effect of pH on the growth of bean plants, and asked learners the questions below, all of which are part of designing an experiment.

(a) Write your research question (or hypothesis).

Table 4: Summary of answers to the above question (n=722)

Frequency in % of answer types	Mark	Code description	Example(s)
3.0	3	Completely correct	<i>What is the effect of pH on the early growth of bean plants?</i>
3.5	2	Answer is a question containing the two variables, but no cause/effect.	<i>What causes the pH to have a negative effect on plants?</i>
3.5	1	Answer is a question which contains one of the variables.	<i>Does it usually affect the plants?</i>
37.7	0	No attempted answer	
24.4	0	Completely wrong	<i>Is that Standerton is because of the people who travels. Get from coal. What is in the universe; star, moon and planets.</i>
0.3	0	Rewrites the question	
27.3	0	Answer does contain one or more words on the topic	<i>Do do like to be hypothesis?</i>

(b) What is the independent variable (the one that you can manipulate)?

Table 5: Summary of answers to the above question (n=722)

Frequency in % of answer types	Mark	Code description	Example(s)
7.9	1	Completely correct	<i>The amount of pH added The amount of acid rain used to water the plants.</i>
41.1	0	No attempted answer	
32.5	0	Completely wrong	<i>What reaction takes place? Cars are independent variable. Gas Is the variable that is standing for itself. AIDS Independent variable is something that not work</i>
0.7	0	Rewrites the question	
3.2	0	Answer is correct if dependent variable had been asked for.	<i>The effect on bean plants</i>
14.4	0	Answer has something to do with the problem	<i>The amount of shelter for the plant so it will not be affected. The health of the plant /soil.</i>

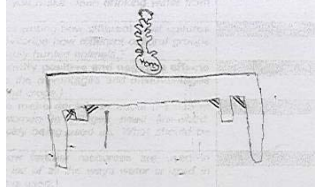
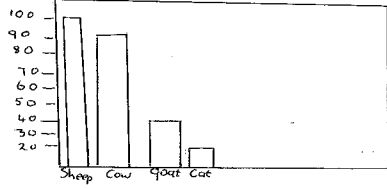
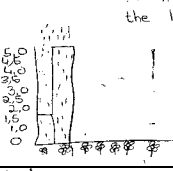
(c) What is the dependent variable?

(d) State two other variables that should be kept constant.

For questions (c) and (d), the data followed much the same pattern as for question (b), so no tables are provided.

(e) Design a table in the space below that you would use to display your results.

Table 6: Summary of answers to the above question (n=722)

Frequency in % of answer types	Mark	Code description	Example(s)																		
0.0	4	Completely correct																			
0.6	3	Close to correct, but with one error	<p>Answer: EFFECT</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>Day</td> <td>Bean plant A</td> <td>Bean plant B</td> </tr> <tr> <td>Monday</td> <td></td> <td></td> </tr> <tr> <td>Tuesday</td> <td></td> <td></td> </tr> <tr> <td>Wednesday</td> <td></td> <td></td> </tr> <tr> <td>Thursday</td> <td></td> <td></td> </tr> <tr> <td>Friday</td> <td></td> <td></td> </tr> </table>	Day	Bean plant A	Bean plant B	Monday			Tuesday			Wednesday			Thursday			Friday		
Day	Bean plant A	Bean plant B																			
Monday																					
Tuesday																					
Wednesday																					
Thursday																					
Friday																					
0.4	2	Draws a labeled table with some appropriate headings	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>Independent variable</td> <td>Dependent variable</td> <td>Control variables</td> </tr> <tr> <td>Bean plant</td> <td>Height of plant</td> <td>Water & sunlight</td> </tr> </table>	Independent variable	Dependent variable	Control variables	Bean plant	Height of plant	Water & sunlight												
Independent variable	Dependent variable	Control variables																			
Bean plant	Height of plant	Water & sunlight																			
0.3	1	Draws a table with at least one appropriate heading																			
53.6	0	No attempted answer																			
25.1	0	Completely wrong, excluding categories below																			
0.1	0	Rewrites the question																			
3.5	0	Draws a furniture type table																			
3.6	0	Draws a graph with no relation to the data																			
1.7	0	Draws a graph with some relation to the data	<p>The more acid rain, the less plants we have ☹</p> 																		
6.1	0	Draws an empty table or a table with no relation to the data, e.g. report card	<p>This term</p> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>MATHS</td> <td>RECORD</td> <td>LOLLI</td> <td>LLRMS</td> </tr> <tr> <td>50/100</td> <td>23/100</td> <td>40/100</td> <td>45/100</td> </tr> <tr> <td>40/100</td> <td>73/100</td> <td>93/100</td> <td></td> </tr> </table>	MATHS	RECORD	LOLLI	LLRMS	50/100	23/100	40/100	45/100	40/100	73/100	93/100							
MATHS	RECORD	LOLLI	LLRMS																		
50/100	23/100	40/100	45/100																		
40/100	73/100	93/100																			
5.0	0	Draws a table with some relation to the data																			

An analysis of the results in terms of socio-economic status (SES)

Explanations for the low levels of achievement could be sought in a number of factors, such as poor language proficiency, poor teaching, poor writing skills, test anxiety, a particular world-view or theoretical framework and socio-economic status of learners. In this research we were interested in understanding learner performance in terms of the socio-economic status of the learners.

At the conclusion of the test, learners were asked to check which of 16 possessions or services they had at home. Examples of items on the list included running water, electricity, telephone, TV and computer. This list is in effect a simple SES index. The index value ranges from 0 to 16, from the simple sum of all items or services a learner has access to at home. Figure 3 represents the percentage of learners in relation to the number of items checked on the list. The greatest cluster of learners appears between 7 and 10.

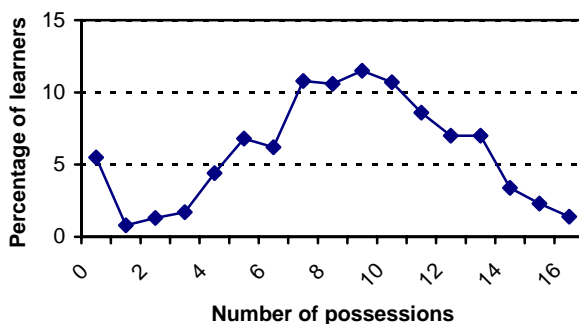


Figure 3: Possessions/services learners have at their homes

The mean scores (expressed as percentages) of learners for the multiple-choice section of the test in respect of the number of possessions/services they indicated are given in Table 7. A positive relationship is evident. Learners who indicated many possessions/services generally achieved high mean scores. The correlation between the number of items and total test score was $r = 0.22$.

Table 7: Relationship between number of possessions/services and mean scores

Number of possessions/services	Number of learners	Mean on test (%)	SE
0	39	23	1.3
1	6	23	4.9
2	9	26	3.1
3	12	21	2.0
4	31	25	1.7
5	48	24	1.3
6	44	29	1.4
7	77	26	1.0
8	75	26	1.1
9	82	24	0.9
10	76	27	1.1
11	61	26	1.2
12	50	26	1.1
13	50	29	1.6
14	24	33	2.7
15	16	37	4.5
16	10	47	7.1

The large-scale PISA study found this same relationship between SES and science learning. The concern raised by Fensham (2003), that there should be "... a sense of equity in relation to the role of the education system as a whole", is valid in South Africa too. The new South African government has spent much energy in equity issues across the board, but success in poverty relief and socio-economic equity has been wanting in rural areas in particular. The widening gap between extreme socio-economic groups may well, from our data, result in an increased science achievement differential in schools and will have to be addressed sensitively with curriculum design.

Discussion

The claim for validity based on the alignment of the test with the intended curriculum is perfectly justified in procedural terms, but nevertheless open to question and debate. It is essential to find ways in which to assess what learners can do and/or should know, particularly in terms of the C2005 outcomes (Malcolm, Kowlas & Gopal, 2004). We hoped that open-ended questions might afford learners the opportunity to express themselves, but they were as poorly answered as multiple-choice questions. Valid assessment of the outcomes on a large scale remains an elusive goal. We would like to contribute to this quest by raising two issues:

- The interpretation of test results when validity is based on the intended curriculum.
- Cultural mismatches of the intended curriculum.

Issues of validity – alignment with policy or practice?

In our study the assessment test was aligned to the intended curriculum (policy) and not the implemented curriculum (practice). The aim of the instrument was to gauge the current performance of a sample of learners in the province on the three learning outcomes of the C2005 Natural Science curriculum, for base-line purposes. There was no attempt to use the results to assess learners as individuals or to predict any future performance. In spite of this argument, the issue of validity is a concern. The extent to which the intended curriculum can be validly assessed by tests depends in large measure on the ability of learners to understand the questions. If the gap between the intended and the implemented curriculum is great, then the validity of an instrument designed to assess the intended curriculum must be suspect in certain respects. The vital question then is how large is this gap?

No direct attempt was made to gauge the extent to which C2005 had been implemented in the 40 schools that comprised the sample. In theory, all schools are implementing C2005 and have made the transition from the content-orientated syllabus of the past to outcomes-based learning. However, the overall progress in Mpumalanga and elsewhere has been uneven, and in general much slower than anticipated (e.g. Ono, Kita & Ndlalane's (2003) description of MSSI in Mpumalanga). In a series of case studies conducted during the previous year (not in the 40 schools of this study), it was found that the gap between the intended and implemented curriculum was large (Rogan, 2004). For example, very few of the learners in the 12 case-study schools engaged in any kind of meaningful hands-on practical work. None had the opportunity to design an experiment as specified by LO1 of the intended curriculum. Hence it might be argued that formal schooling has hardly equipped learners to even make meaning of the questions geared to this outcome, let alone answer them. Hence the validity of the assessment of this intended outcome remains problematic.

On the other hand, the results provide some evidence that the gap is being partly bridged. The Grade 9 learners performed significantly better than those in Grade 8 overall and in Outcomes 1 and 3, suggesting that learning that is measurable is occurring. Outcomes 1 and 3 have introduced outcomes that have not been dealt with in the past, whereas Outcome 2 deals with knowledge. It is therefore perhaps significant that the grade 9 learners, who had been exposed

to C2005 for an additional year, did better in these two outcomes. Furthermore, the increase in performance with SES, which is a world-wide phenomena, also points to some measure of validity. Provided the instrument is well-constructed (and evidence to this effect is provided in the following article by Boone and Rogan), then as the gap closes, so the validity improves. If MSSI and other efforts by the Mpumalanga Department of Education to improve the implementation are successful, then not only should the results based on this test improve with time, but they can also be interpreted with more confidence.

No consideration of validity in the context of this study can ignore the question of language. (See for example, Howie, 2002, 2003). The answers to the open-ended questions raise doubts about the ability of many learners to understand the question and/ or express themselves in English, a second language for almost all of them. This lack of ability compromises the validity of the test. Although open-ended questions potentially have the ability to assess what learners know rather than what they do not know, such potential depends on language proficiency. Validity then depends in turn on the test takers' proficiency in the language of the test. As long as tests are given in languages in which the learners are not proficient, validity is compromised.

Alignment with cultural values

The fact that outcomes are apparently not being achieved may reflect an under-explored, but historic, tension. The Natural Science policy document (Department of Education, 2002) specifically encourages the inclusion of Indigenous Knowledge Systems (IKS).

Indigenous or traditional technologies and practice in South Africa were not just ways of working; they were ways of knowing and thinking. Traditional technologies and practices often reflect the wisdom of people who have lived a long time in one place and have a great deal of knowledge about their environment. Wisdom means that they can predict the long-term results of decisions, and that they can recognise ideas which offer only short-term benefits (p. 10).

Yet at the same time, traditional/indigenous forms of knowledge transmission are being devalued. One aspect of C2005, which appears to have become part of our new vocabulary is teacher as facilitator. By this it is understood that teachers should give up their role as the presenter of information and help the learners to acquire knowledge through interaction with curriculum materials and one another. The value of this method of learning is not in dispute. But it is not the only valid way of structuring learning. Africa has a long and still active tradition of oral learning, in which the art of rhetoric is greatly valued, and many teachers excel at it. It is a means of teaching that makes sense in schools with large classes and meager resources. Learners respond to and engage in skilful rhetoric, as we have observed in some of our case studies (Rogan, 2004). C2005 can build on current strengths rather than undermine them; the outcomes of C2005 can be achieved in a variety of ways.

This cultural tension is not new. The curriculum revolutions of the 60's did not bypass Africa. However, much of the curriculum work in Africa consisted of adapting first world innovations. This, despite the fact that, "*even before independence most colonial powers had become aware that the wholesale transplanted of their system of education did not find a fertile soil in Africa.*" (Ogunniyi, 1996, p. 70). Curricula are embedded with cultural values. For example, a curriculum designed in a Western country may place value on individual development, which may well be at odds with the African notion of ubuntu. Similarly, some manifestations of child-centred learning advocated in Western culture – such as encouraging a child to question authority – do not sit well in African cultures. As de Feiter and Ncube (1999) point out,

It presents pupil-centred education as the ultimate form of good education in line with the individualistic Western culture. In fact, this may not necessarily be the best teaching strategy in all cultures (p. 89).

It is possible that one of the reasons that C2005 has not delivered its intended outcomes is that its underlying philosophy of teaching and learning are not aligned with dominant cultural values.

Conclusion

The data provided in this paper provide a description of Grade 8 and 9 learners' performances in outcomes pertaining to the new Natural Science curriculum. The research is to provide a baseline assessment of learner performance, as part of assessment of the impact of the MSSSI project in the Mpumalanga province. The tests indicate low levels of performance in each of the three Learning Outcomes of the Natural Science curriculum, in spite of C2005 implementation and the MSSSI. The levels of performance correlate with socio-economic status, whereby the higher performing students generally rate higher on SES measures.

Reflecting on the performance data, we raise a number of issues. One concerns the validity of tests, in the face of, for example, language differences. The second concerns curriculum alignment: (a) the alignment of curriculum policy and practice, (b) the alignment of curriculum policy and cultural values. We argued that learner performance will be little affected in relation to the intended policy if what actually happens in classrooms is poorly aligned with the intended policy. This alignment issue needs to be considered when designing instruments for measuring learner achievement. Secondly, we suggest that measuring instruments and ultimately educational philosophies need to be sensitive to prevailing value systems in the schools, if the intended and practiced curricula are to support one another. This has ramifications for assessment and testing.

References

- De Feiter, L.P. & Ncube, K. (1999). Toward a comprehensive strategy for science curriculum reform and teacher development in Southern Africa. In Ware (ed.) *Science and Environment Education: Views from Developing Countries*. Washington, DC: The World Bank.
- Department of Education (1997). Curriculum 2005. Retrieved on 07/05/2004 from <http://www.polity.org.za/govdocs/misc/curr2005.html>
- Department of Education (2002). Revised National Curriculum Statements Grades R-9: Natural Sciences.
- Fensham, P.J. (2003). *PISA science: Pointing the way forward for school science*. Paper presented at the annual SAARMSTE conference, Swaziland, January 2003.
- Fullan, M. (1999). *Change forces: The sequel*. London: Falmer.
- Howie, S.J. (2002). *English language proficiency and contextual factors influencing mathematics achievement of South African secondary school pupils*. Enschede: Doctoral dissertation, University of Twente.
- Howie, S.J. (2003). Language and other background factors affecting secondary pupils' performance in Mathematics in South Africa. *African Journal of Research in Science, Mathematics and Technology Education*, 7, 1-20.
- Lawrenz, F., Huffman, D. & Welch, W. (2001). The science achievement of various subgroups on alternative assessment formats. *Science Education*, 85, 279-290.
- MacBeath, J. (1999). *Schools must speak for themselves: The case for school self-evaluation*. London: Routledge.

MacBeath, J., Schratz, M., Jakobsen, L. & Meuret, D. (2000). *Self-evaluation in European schools*. London: Routledge Falmer.

Malcolm, C., Kowlas, L. & Gopal, N. (2004). *Assessing to find out what children do know*. In: Buffler, A. & Laugsch, R.C. (eds). Proceedings of the 12th annual conference of the Southern African Association for Research in Mathematics, Science and Technology, Cape Town, South Africa, 598-602.

Ogunniyi, M. B. (1996). Improving Science and Mathematics Curriculum in African Schools: A Synopsis. In C. Stoll, L. de Feiter, H. Vonk, & J. van den Akker (eds), *Improving Science and Mathematics Teaching in Southern Africa: Effectiveness of Interventions*. Amsterdam: VU University Press.

Ono, Y., Kita, M. & Ndlalane, T. (2003). *Peer teacher learning: Introduction of lesson study (Jyugyo Kenkyu) as school-based INSET in South Africa*. Paper presented at the ICET conference, Melbourne, Australia.

Rogan, J.M. (2004). Out of the frying pan ... ? The implementation of Curriculum 2005. *African Journal of Research in Mathematics, Science and Technology Education*, 8(2), 165-179.

Shavelson, R.J., McDonnell, L.M. & Oakes, J. (1987). *Indicators for monitoring mathematics and science education: A sourcebook*. Santa Monica, California: The RAND Corporation.

Spady, W.G. (1994). *Outcomes-based education: Critical issues and answers*. Arlington: American Association of School Administrators.