CENTRE FOR SOCIAL SCIENCE RESEARCH

Social Surveys Unit

INDICATORS OF PERFORMANCE IN SOUTH AFRICA'S PUBLIC SCHOOL SYSTEM

Jeremy Seekings

CSSR Working Paper No. 10

September 2002

ISBN No.: 0-7992-2130-9

Centre for Social Science Research
University of Cape Town
Private Bag
Rondebosch
7701
Telephone: (021) 650 4656

Website: http://www.uct.ac.za/depts/cssr

Indicators of Performance in South Africa's Public School System

Education is the largest area of expenditure in the South African state budget. In the 2001/02 financial year, nearly R60 billion, or 22% of the total budget, was notionally allocated to education. The bulk of this budget – R44 billion in 2001/02 – is transferred to the nine provincial education departments for primary and secondary schooling. In comparison with other middle-income countries, the government's expenditure on education, relative to both the total government budget and the GDP, is high.

Yet the quality of schooling in South Africa is believed to be relatively poor. Pass rates in the matriculation examination are low, and studies comparing literacy and numeracy among students in different countries almost routinely put South Africa at the bottom of the list. The important 1999 study, *Getting Learning Right* (Taylor and Vinjevold, 1999), painted a sorry picture of the state of South African schooling. The national Department of Education itself, in its own documentation, acknowledges the many weaknesses of the schooling system (see, for example, Department of Education, 2000a).

The result is growing concern over the mismatch between expenditures and outcomes. This is emphasised most often in documents from the Department of Finance. The 2000 *Interdepartmental Fiscal Review*, for example, notes that 'at almost 6% of GDP, South Africa has one of the highest rates of government investment in education in the world, yet education outcomes lag behind those in comparable countries' (Department of Finance, 2000: 29). Similar concerns are raised by institutions such as the Financial and Fiscal Commission and the Auditor-General. They are echoed, at least, in documents from the national Department of Education. In 2000 Parliament's Standing Committee on Public Accounts (SCOPA), exercising its mandate to monitor public expenditure, also commissioned a study of the efficiency of public expenditure on schooling.

South Africa's problem with 'value for money' may be extreme, but it is not entirely unique. In the USA, '…the costs of education have been growing far more quickly than the benefits'.

> 'After allowing for inflation, the amount spent on each pupil in America's schools has increased unabated for a century... During the past three decades, however, student performance has, at best, stayed constant and may have fallen…' (Hanushek *et al*, 1994: xix).

Other developing countries are also worried about the inefficient use of scarce public revenues ('inputs') in producing educational 'outcomes' in terms of students' performance. As one recent study writes:

> 'Most developing countries now recognise that investing in education, particularly primary education, provides an essential bedrock for economic and social development. In past decades governments emphasised expanding enrolment, but as coverage rose, the problems of low completion rates and inadequate student learning came to the fore. Policymakers need information on the costs as well as on the impact of different methods of improving schooling outcomes' (Tan, Lane and Lassibille, 1999:493).
>
> 'Unfortunately', as this study itself continues, 'the literature on the quantitative relationship between inputs and outcomes in education is sparse, and most

> developing countries have only a nascent capacity to conduct their own context-specific research and evaluation' (*ibid*).

The relationship between expenditures and outcomes is important not because of some ideological concern with 'efficiency' or because of any supposed need to reduce expenditures. Rather, its primary importance lies in helping education policy- or decision-makers to identify the most effective ways of producing the best possible education for children given the inevitably limited resources. Is the education of school students served better if public resources are allocated to reduced class sizes (requiring extra classrooms and especially extra teachers), or to better trained teachers, or to the supply of textbooks and other 'learning support materials'? How should schools be managed or organised so as to achieve the best possible education for their students? To answer such questions, two things are required. First, data needs to be collected to allow for careful empirical analysis of schooling. Secondly, 'experiments' need to be run with alternative ways of investing resources and managing schools. This working paper examines one half of the first of these topics: the indicators of educational outcomes.

No assessment of efficiency can be undertaken without accurate data on expenditures and reliable and meaningful indicators of desired educational outcomes. Until recently, there was little recognition in South Africa of the importance of either of these. Crouch wrote, as recently as 1999:

> 'What one normally thinks of as 'education indicators' have not been widely used in South Africa's past or, when used, were often used without sufficient critical understanding of their purposes or limitations. The tradition is hard to break. With a few (key, but often misused) exceptions, such as the Std. 10 pass rates, per learner expenditures and the learner-educator ratio, the system proceeds largely unaided by the use of indicators.'

Crouch muses on the effect of this on education management:

> 'The experience must be something like that of flying a Boeing 747 with, say, only a fuel gauge and a weight gauge prior to take-off: no altimeter, no speed gauge, no radar, no engine temperature gauges, no compasses, no gyroscopes, no oil pressure gauges, etc. Exhilarating, but prone to surprises. Not an enviable position for the Minister and the MECs [i.e. provincial Members of Executive Councils in charge of education] to be in – one expects that if they want roller-coaster rides they would rather get them at an amusement park than in the sector they are accountable for' (Crouch, 1999).

There seems to be an attitude of hostility when it comes to examination of costs (which must surely in part reflect simple self-interest among education bureaucrats; there is also hostility to the use of quantitative analysis (perhaps born out of the anti-positivist stance of South African social sciences for many years – see Muller (1999)), as well as to formal assessment of outcomes.

The Department of Education itself recognises that '…there is no comprehensive data on the quality of teaching and learning, or on the educational standards achieved in the system' (Department of Education, 2000b: 3). The introduction of indicators of outcomes (i.e. the outputs of the education system) has been and remains *especially* controversial. An emphasis on *outcomes* themselves is not controversial: 'outcomes-based education' (OBE) emphasised outcomes just as much as alternative educational philosophies and systems of curriculum design. What is controversial is the way in which outcomes are assessed or tested. Formal testing is very different to informal assessment, and Curriculum 2005 seems to have been understood as entailing informal

assessments. There are at least three lines of criticism regarding more formal systems of assessment: available tests focus on too restricted a range of skills; education can suffer if teachers simply 'teach to the test'; and assessment should focus on progress or value-added, rather than simply levels attained.

Among some educationalists there appears to be a general suspicion of testing. On the other hand, there are serious, perhaps insuperable, practical problems with complex forms of assessment. One of the most telling criticisms of the Department of Education's adopted OBE/Curriculum 2005 was that few teachers were able to conduct adequate assessments of the range of outcomes identified in policy guidelines.

Hostility towards the assessment of student achievement was evident in the Department of Education's April 2001 Colloquium on systemic evaluation (see below). In one group discussion at the colloquium, participants were divided over the importance of student achievement: 'On the one hand, some members accept that learner achievement will always be a key indicator of the state of education; on the other hand, the feeling is that it would be dangerous to over-emphasise this one indicator and ignore all the factors which influence learner achievement' (Department of Education, 2001b: 22). This reveals confusion over means and objectives. The fundamental objective of public investment in education is to produce higher levels of student achievement. The most important indicators of all are those that demonstrate how far students are achieving the desired performance goals. Most other indicators focus on aspects of the process of achieving these goals, not on whether the goals themselves are being attained. Thus, indicators of teacher training, or class sizes, or of physical infrastructure, are actually at most indicators of 'intermediate variables', not final outputs. Indeed, it probably makes more sense to think of these as 'inputs' into education, in that financial investments convert fairly readily into teacher training, class sizes and physical infrastructure. The difficult questions concern the relationship between these inputs and student achievement.

Hanushek *et al* (1994) note a similar reluctance to evaluate systematically the performance of students in the USA:

> 'Two reasons underlie this reluctance to judge. The first is that defining success in schooling is hard. Intuitively, school personnel have long resisted standardised testing and assessment. Instead of using a fixed set of criteria to evaluate students, educators have preferred more flexible systems that allow each student to emphasise his or her own achievements. Yet flexible measurement quickly degenerates into a collection of semi-independent observations and impressions that cannot be compared. The second difficulty is assessing schools is establishing responsibility and developing appropriate accountability. Even if some measurement of performance is accepted, using it to assign praise or blame is difficult because so little is understood of the inner workings of the education process. Schools, as we have repeatedly stressed, are just one part of education, and schools themselves are complicated institutions involving many components' (Hanushek *et al*, 1994: 25-6).

When students have been assessed, it has generally been in order to indicate to teachers (and sometimes parents) how *individual* students have been progressing, and what their individual needs might be. Such individual assessments also provide pressures on or incentives to students to study harder. Assessment has rarely been conducted to generate information on the performance of different kinds of educational policy or on the efficacy of using resources in particular ways.

The formal policies of the Department of Education recognise the need for indicators (as we shall see in detail below). But it is probably true to say that this recognition has been slow in coming, and there remain strong pockets of ambivalence within the education system. More important, there is little clarity on how indicators should be *used*. At the outset, therefore, it is necessary to sketch the *roles* that indicators can play. Later, we shall examine whether the indicators being developed in South Africa are being used to fulfil these roles fully.

The Roles of Performance Indicators

There is a healthy debate on the value of different kinds of indicator of educational outcomes. Taylor (1999) provides a useful introduction to this, whilst Crouch (1999) and Gilmour (2000) provide valuable analysis and insights.

Taylor begins by distinguishing between three different kinds of assessment, noting how they play different roles.

- Public exams are primarily 'summative' in that they assess the achievement of pupils for the purposes of certification. Public examinations may also play a 'formative' role, in that they inform policy-making, but this is secondary.
- National assessments are primarily formative, in that their purpose is to measure the performance of the education system with a view to improving that performance.
- Classroom-level assessment is primarily 'diagnostic' and formative in that it is intended to identify the needs of individual students in order to inform teaching practices. Insofar as this kind of assessment informs progression into the next grade, it performs also a summative role.

Each of these ways of assessment will require different indicators. It is likely that the kinds of indicators which are most useful in terms of a student's own progress would be prohibitively expensive to run on a countrywide basis, or could only be run if schools were already operating at very high standards – in which case the need for indicators for national planning would be reduced. In the United Kingdom, Taylor points out, the assessment scheme introduced in the early 1990s sought to combine national assessment with classroom-level assessment; the scheme proved unworkable, because it placed excessive demands on teachers (Taylor, 1999:187). This lesson is very relevant to South Africa, where one powerful criticism made of the Outcomes-Based Education initiative was that teachers lacked the skills to assess outcomes.

The European Union suggests a set of sixteen quality indicators. These divide into four categories:

- attainment: in (1) mathematics, (2) reading, (3) science, (4) information and communication technology, (5) foreign language, (6) 'learning to learn' and (7) civics;
- success and transition: (8) drop-out, (9) completion of upper secondary education, (10) participation in tertiary education;
- monitoring of school education: (11) 'evaluation and steering of school education', (12) parental participation;
- resources and structures: (13) education and training of teachers, (14) participation in pre-primary education, (15) number of students per computer and (16) educational expenditure per student (cited in Baird, 2001).

Use of this list would serve to generate internationally-comparable data on the performance of a country's education system, and would therefore inform policy-making. However, detailed policy-making requires a wider range of data.

Crucially, if a country needs to know precisely how it should be investing its education budget, it needs to be able to trace the success of different investment decisions – i.e. different educational interventions – in terms of student achievement.

Matriculation Results

The results of the Matriculation (or Senior Certificate) examination, sat at the end of Standard 10 (i.e. Grade 12), are widely regarded as the definitive measure of quality in the schooling system.[1] Pass rates for the five years between 1996 and 2000 are presented in Table 1. The Table shows that results were poor in 1997, 1998 and 1999, but improved markedly in 2000. There are big variations between provinces. The worst results in this five-year period was the Northern Province's 32% in 1997. The best result was the Western Cape's 81% in 2000.

Table 1: Matric Results by Province (Percentage Pass Rates)

	1996	**1997**	**1998**	**1999**	**2000**
Eastern Cape	49	46	45	40	50
Free State	51	43	43	42	53
Gauteng	58	52	56	57	68
KwaZulu-Natal	62	54	50	51	51
Mpumalanga	47	46	53	48	53
Northern Cape	74	64	65	64	71
Northern Province	39	32	35	38	51
North West	70	50	55	52	58
Western Cape	80	76	79	79	81
Average	54	47	49	49	58

Source: RSA, 2000: Table 2.7; 2000 results from the *Cape Argus* 29 Dec 2000.

The big improvement in 2000 was in Northern Province, where the pass rate rose from 38 to 51% Kader Asmal attributed the improvement to the completion of the curriculum and the establishment of weekend schools in some subjects.[2] The 2000 exams certainly saw some improvements in key subjects. Higher grade passes in mathematics rose from 54 to 65% and in physical science from 61 to 65%.[3]

Performance varies considerably between subjects. The 1998 results for selected subjects are set out in Table 2 below. Whilst pass rates are high for subjects like English and Afrikaans, they remain very low in others (notably mathematics), especially at the higher grade.

[1] For a critical summary of what it is that the matric exam. actually tests, see Taylor (1999: 200-202).
[2] *Sunday Independent*, 31 December 2000.
[3] *Cape Argus*, 29 December 2000.

Table 2: Matric Pass Rates by Subject, 1998

Matric subject	**matric candidates sitting the exam**	**pass rate %**	**pass rate on higher grade %**
Mathematics	279 702	42	7
Physical Science	157 174	66	17
English	562 156	89	65
Biology	432 295	57	13
History	178 846	58	15
Afrikaans	464 049	89	63

Source: SAIRR, *South African Survey 1999/2000*, p.117. NB these figures include students sitting for less than six matric subjects.

On the matriculation examinations, we have a detailed analysis of the 2000 results, undertaken by the national Department of Education (2000d). We also have Luis Crouch's valuable scrutiny of the use of matric pass rates as an indicator of system quality (1999). He concludes that they play some roles better than others: they provide a better indicator of the overall performance of the schooling system than they do of the performance of individual schools. The reason for this latter weakness is that 'they measure too many things other than the quality of individual schools' inputs and production processes' (*ibid*: 28). One of the major confounding factors is the family and community background of the pupils at the school. As Crouch (together with Mabogoane, 2001) has shown, as much as 25% of the performance gap between schools may be due to the social and economic inequality of students' backgrounds. If we want to use matriculation pass rates as an indicator of the performance of individual schools, we need to take into account inequality in the background of pupils and other factors outside of schools' control. In other words, we want to estimate the 'value added' in the secondary schools, not just the final level achieved. Crouch and Mabogoane have done some crude estimates of this for Gauteng. They show that the top 10% of schools in terms of straight pass rates have an average pass rate of 99.9% whereas their socio-economic profiles would lead us to expect a pass rate of about 88%. These schools are indeed 'over-performers', in that they perform better than might be expected using crude data. But they are certainly not the top *over*-performers. The top 10% of schools defined in terms of the level of over-performance, i.e. pass rates adjusted for socio-economic background, over-performed by an average of 30%. Many of these schools were in poor areas, but nonetheless achieved good (and sometimes very good) results. It must also be noted that Crouch and Mabogoane's analysis of the use of matric pass rates as an indicator was written in 1999, prior to the controversy over the 2000 results and (specifically) how the higher pass *rate* should be interpreted.

Matriculation results need to be treated with some caution, however. Consider three caveats regarding the apparently improved results in 2000. Firstly, repeaters were discouraged (or excluded?) for the first time in 2000. The number of repeating students was reduced by 90% between 1999 and 2000, from nearly 63 000 to just 6 300.[4] The teachers' union SADTU wondered whether these non-repeaters had simply dropped out, joining the ranks of the unemployed.[5]

Secondly, continuous assessment coursework marks were included for the first time in 2000 for some subjects in seven of the nine provinces; these marks counted for up to 25% of the final 'examination' result[6] (although another report said that was in only three provinces[7]).

[4] *Sunday Times*, 31 December 2000.
[5] *Cape Argus*, 29 December 2000.
[6] *Sunday Times*, 31 December 2000.

Educationalists disputed the significance of this reform. Some were reported as saying that it could account for an overall increase of as much as 5% but Minister of Education Asmal said it could account for 'about one and a half per cent' only.[8]
Thirdly, in September 2000, the Department of Education reportedly asked schools to encourage students to drop from higher grade to standard grade if there prospects of passing at the higher grade were poor.[9]

In addition, there remains a suspicion that the results are manipulated. This was proved to be justified in the case of Mpumalanga in 1998. In December 2000, the spectre of manipulation was raised anew. In a letter to the press, a 'concerned teacher' in the Western Cape claimed to have been told that 'matric results are manipulated, based on the socio-economic conditions of a particular school's environment'. The results of students attending 'good' schools are not changed, but the results of students in 'disadvantaged' schools are adjusted upwards. The letter-writer also alleged that internal examination results within schools were adjusted upwards under pressure from school inspectors.[10] The leap in the pass rate in 2000 also raised some eyebrows. Educationalists had expected some improvement, in part because there had been few strikes by teachers. But, on the eve of the announcement of the 2000 results, Jonathan Jensen warned that, 'If we see jumps in the order of between 5% and 8%, serious questions should be raised about whether there has been fiddling'.[11] The overall improvement was nine percentage points. But Asmal dismissed Jensen's criticism as 'a guttersnipe argument'.[12]

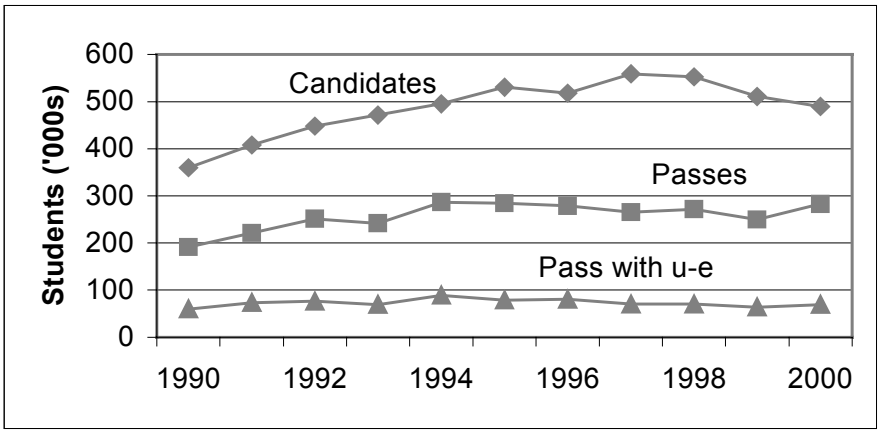
Some adjustments to marks are made openly. Marks are 'standardised' in each province, under the supervision of the South African Certification Council. Marks in any examination are compared to 'a five-year norm' and adjusted to result in 'a normal pass rate' (Hartley *et al*, 1998a: 5). The Certification Council is responsible for certifying the process as well as the results.

Perhaps the most important qualification of the matric exam results, however, is the relationship between pass rates and the number of students writing the exams. Higher pass rates might be because fewer students, and especially fewer weak students, are attempting the examination. Figure 1 shows the aggregate number of students attempting and passing the matriculation examination (with a separate line for passes with 'university exemption ('u-e' in Figure 1).

---

[7] *Sunday Independent*, 31 December 2000.
[8] *Ibid.*
[9] *Business Day*, 28 December 2000.
[10] *Cape Argus*, 29 December 2000.
[11] *Business Day*, 28 December 2000.
[12] *Sunday Independent*, 31 December 2000.

Figure 1: Numbers of Students Sitting and Passing the Matric Exam 1990-2000



The number of students passing the matriculation examination can also be compared usefully with the number of students in the age cohort, and with the number entering the schooling system twelve years or so previously. In a perfect schooling system, in which all pupils passed each year and proceeded to the next grade, the number of students in Grade 1 would be the same as the number passing the matric examination twelve years later. Table 3, prepared by the Research Institute for Education Planning (RIEP) at the University of the Free State, presents data on the flow of pupils through the school system.

Table 3. Flow of Pupils Through the School System, 1983-1994

	All pupils		White and Indian pupils	
	number	% of 1983 grade 1	number	% of 1983 grade 1
Grade 1 (1983)	1 126 050	100	104 184	100
Grade 10 (1992)	491 378	44	96 186	92
Grade 12 (1994)	464 524	41	81 632	78
Grade 12 passes (1994)	212 541	19	74 802	72

Source: Calculated from data from RIEP, reproduced in FFC (1998): 58.

The total number of pupils in Grade 10 in 1992 and Grade 12 in 1994 was only 48% and 44% of the number in Grade 1 in 1983. The figures for white and Indian students were very different from the figures for African and coloured students. Whereas the number of white and Indian students in Grade 12 in 1994 was 78% of the number in Grade 1 twelve years previously, the equivalent proportion for African students was just 43% and for coloured students 26%.

The figures in Table 3 are probably almost meaningless, however. Both the Grade 1 and Grade 12 numbers are inflated by over-enrolment. As Crouch and Mabogoane put it, the ensuing ratio is 'relatively useless' (1997: 17). They calculate that 'the total flow-through in the system is much higher than is habitually believed' (*ibid*: 25):

> 'It has become common wisdom that only about 30%–35% or so of African youth make it through the 12-year system as matric-level enrolees. We believe the number

> is much higher, depending on methodology and data sources. However, it is true that only about 25%–30% of each cohort of young adults (of all groups) is entering adult life with a matriculation pass.'

As they acknowledge, however, many – perhaps half the number - of children do fall by the wayside before reaching Grade 12. A further indication of this is the proportion of young men and women who have not reached or completed Grade 12. Representative surveys (such as the October Household Survey) tell us that in no age cohort have more than 40% of the population completed Grade 12 (Lam, 1999: 36).

Enrolment, Progression and Repetition Rates

Among the major problems with using matric pass rates as an indicator of performance are, first, that many pupils do not get as far as Grade 12, and, secondly, the pass rate does not indicate how many years the relatively successful pupils took in getting to Grade 12.

A useful summary statistic indicating throughput is the number of school years provided for pupils in each age cohort per matric exam pass at the end of Grade 12 (Standard 10) by a student in that cohort. If all students progressed into the next grade each year, and none dropped out, and all students taking the matric exam passed it, then the average number of school years per matric pass would be twelve. However, because many students drop out from, fail, or repeat grades, and many students do not make it to Standard 10,yet alone pass matric, the number of school-years per matric pass is much higher than this. Work by Research Institute for Education Planning at the University of the Free State (RIEP) suggests that each matric pass by a black student in 1981 required forty-five student years. In 1983 and 1984 the ratio narrowed to thirty-nine years (cited in Crouch and Mabogoane, 1997: 18). Crouch and Mabogoane use a variety of techniques to calculate data for more recent years. They conclude that 'the system "invests" some eighteen years of learner effort to "produce" one Grade 12 enrolee, and some 36 learner/years of effort to produce one Std 10 pass' (*ibid*: 25). They add that there are big variations between provinces.

> 'In some provinces those sitting for the exams are already a select elite, while in others learners flow through as fail at the last point. This renders provincial comparisons very difficult' (*idem*).

The 'real tragedy', they remind us, is that young people are 'investing more than twelve years of their lives, to achieve only a 1:4 or 1:3 chance of entering adulthood with a Std 10 pass' (*idem*). It is unclear how this statistic for South Africa compares with other countries, but it is clearly worrying.

To put the matric pass rate data into perspective, we need additional data on enrolment, repetition, and failure and drop-out rates. Unfortunately, our data on all of these are flawed.

Enrolment

South Africa has high enrolment rates in comparison with many other developing countries: in the mid-1990s, only about 15% of eighteen-year-olds in South Africa were not enrolled in school. This is generally because they don't want to be there and not because there are no places in local schools (Case and Deaton, 1999: 1058; Crouch and Mabogoane, 1997: 18). In South Africa in 1995, the mean number of years of completed schooling for people between the ages of 20 and 24 years was 9.5. In Brazil, the mean was just 6.65 years. There was also much less inequality in years of completed schooling in South Africa than in Brazil (Lam, 1999). In other words,

enrolment rates are high even among children from poor households in South Africa. South Africa had universal schooling up to the age of fifteen before it was made compulsory (see also Anderson *et al*, 2001).

One measure of enrolment is the 'gross enrolment ratio', i.e. the ratio of pupils enrolled in primary and secondary schools to the total number of seven-to-eighteen-year-olds in the population. The most recent data from the Department of Education suggests that the gross enrolment ratio for South Africa as a whole is 100. In several provinces, most notably Eastern Cape, KwaZulu-Natal and Mpumalanga, it is above 100%, indicating the presence in school of under- or over-age students (Department of Education, 2001a: 2).

There are some problems with the data, however. The results of the most recent, 1996, Census were inconsistent with data from sample surveys and the Department of Education's own enrolment data. Whereas the latter suggested an aggregate enrolment rate of about 74% among 5-24 year-olds, and over 95% in the 7-15 age group, the 1996 Census suggested an enrolment rate of just 64% for 5-24 year-olds; the Census data does not permit the calculation of an enrolment rate for the 7-15 age group (Crouch, 1998). The Department of Education recognises that their own Education Management Information System (EMIS) data is likely to exaggerate the number of students in the system, as principals have powerful incentives to exaggerate the number of students in their schools. The Department of Education's own reports, however, indicate that there has been only one major attempt to identify ghost students (as well as ghost teachers), in KwaZulu-Natal in the late 1990s. There is a need for more accurate data on enrolment, drop-out and repetition rates, as well as the promotion of under-performing pupils.

In South Africa, as Crouch has emphasised, the problem is one of 'over-enrolment' in some grades.

Repetition Rates

Many pupils repeat grades more than once. Data from RIEP suggests that repetition rates among black pupils in the early 1990s decline from about 18% in Grade 1 to 12% in Grade 7, then rise sharply into the 20s in Grades 8 to 11, and peak at 40% in Grade 12. Drop-out rates run at about 19% in Grade 1, less than 10% in Grades 2 to 7, and above 10% in Grades 8 to 11 (cited in Crouch and Mabogoane, 1997: 23). Crouch and Mabogoane suggest that these repetition rates are probably underestimated, and the dropout rates overestimated, especially in the lower grades. They suggest that probably only 60% of the six-year-old black children in Grade 1 in any year progress to Grade 2 the following year, with 40% repeating Grade 1. By the third year, they estimate, fewer than half have reached Grade 3, whilst 16% are still in Grade 1 and 33% are in Grade 2 (24). Repetition is most common in Grades 1, 8 and 12.

In 1999 data was collected on the number of years that a large sample of pupils had spent in each grade. Nationally, 16,9% of pupils reportedly spent two years in Grade 1, and 3,1% spent three years in Grade 1. Grade 1 repetition rates were highest in North-West, where a massive 22,1% of pupils had spent three years in Grade 1, and lowest in the Western Cape and Gauteng, where fewer than 10% of pupils spent more than one year in Grade 1 (Department of Education, 2000a: 37). Such survey-based data needs to be treated with some caution, however. If it is collected from pupils themselves, it is sometimes the case that pupils are unable to account for all the years that they spent in or out of school, i.e. they underestimate repetition rates.

In a case-study of one high school in a Cape Town township, Anderson *et al* (2000) found that students had repeated on average two grades, mostly in primary school. Few students repeated grades in secondary school. Anderson *et al* found significant rates of social promotion

of students who had failed in Grades 8 and 9. Indeed, most students who failed in Grades 8 or 9 were promoted and did not repeat.

In the absence of public examinations at multiple levels, it is unwise to read too much into 'failure' and 'repetition' rates in lower grades. Crouch conducted a careful analysis of repetition rates, finding that repetition rates in primary schools were highest among students who began their formal schooling at an early age. In response to this the national Department of Education introduced a minimum age for admission into Grade 1. Schools were instructed at the beginning of 2000 to exclude children who would not turn seven during the school year. This prevented parents from using primary school as a free creche facility for young children. The number of students in Grade 1 in Gauteng dropped by one-third. Porteus (2000: 21, fn 39) points out that early childhood educationalists were divided over the educational effects, with some arguing that using schools as a childcare facility was preferable in educational terms to leaving children alone or even preventing a parent from working. In 2001 it appears that repetition rates have declined. It should also be remembered that studies using numeracy and literacy tests (such as TIMSS, see below) all found huge variations in literacy and numeracy levels among students in the same grades.

The 1998 Assessment Policy provides clear guidelines for progression.

> 'It is expected that in the main, learners will progress with their age cohort. Where it is felt that a learner needs more or less time to demonstrate achievement, decisions shall be made based on the advice of the relevant role-players: educators, learners, parents and education support services [i.e. including educational psychologists]. If a learner needs more time to achieve particular outcomes, he or she need not be retained in a grade for a whole year. No learner should stay in the same [three-grade] phase for longer than four years, unless the provincial Head of Department has given approval based on specific circumstances and professional advice' (Department of Education, 1998: 14, para.32)

In the absence of clear criteria to measure student achievement, it is impossible to assess the significance of progression and repetition rates. Reducing repetition through either exclusion or social promotion may cut costs but it is likely to worsen performance in terms of pupil achievement.

The Department of Education's New Assessment Policies

The national Department of Education is supposedly committed to the development of new forms of 'quality assurance'. The Minister of Education is required under the National Education Policy Act (1996, Section 8) to monitor and evaluate standards of education provision, delivery and performance. A Directorate of Quality Assurance was established, and Quality Assurance (QA) structures were established in the provincial departments. A National Quality Assurance Co-ordinating Committee was established to provide for liaison between national and provincial officials.

The mandate of the QA has been described as follows:

> 'It has the responsibility for providing timely and accurate information about the achievement of South African schools and learners. Therefore, we have responsibility for large-scale assessments; quality indicators; evaluation of schools and co-ordination of our involvement in international studies. The mandate is broad and it

> seeks to bring all the education accountability and quality initiatives under one domain'(Department of Education, 2001b: 10)

The proposed new forms of quality assurance include:

- the assessment of individual students through continuous assessment and public examinations, prior to Grade 12 (in terms of national Assessment Policy gazetted in December 1998);
- 'systemic evaluation' at Grades 3, 6 and 9 (proposed in the same 1998 Assessment Policy); and
- 'whole school evaluation' (in terms of a national policy published in 2000).

The Department of Education formally recognises that the development of new indicators of educational outcomes are integral to all of these forms of quality assurance. In practice, however, the national Department appears to be dragging its feet on the development of indicators.

NGOs including the Joint Education Trust (JET) were involved in the development of tests to be used in systemic evaluation in Grades 3 and 6, but the national Department of Education (or key officials therein) opted instead to develop an entirely new set of tests through the parastatal Human Sciences Research Council (HSRC), which were supposedly more compatible with OBE. The JET tests have been used widely in different regions. The alternative HSRC tests were piloted at the end of 2000. Whether there is any significant difference between the two sets of tests is unclear.

The Proposed New Assessment Policy

In December 1998 the Department of Education (1998) gazetted a new Assessment Policy covering schooling up to and including Grade 9. The policy entailed the identification of 'expected levels of performance' for each grade ('and phase'), new methods of assessment, and new reporting requirements. Together, the Department claimed, these would allow schools, students and parents to assess accurately the progress being made by individual students. Students' progression through grades would take into account their achievement in terms of the published 'expected levels of performance'.

The Department's definition of assessment pointed to the multiple roles of the new policy:

> 'Assessment is the process of identifying, gathering and interpreting information about a learner's achievement, as measured against nationally agreed outcomes for a particular phase of learning. It involves four steps: generating and collecting evidence of achievement, evaluating this evidence against the outcomes, recording the findings of this evaluation and using this information to assist the learner's development and improve the process of learning and teaching' (*ibid*: 9).

The details of the Assessment Policy indicated, however, that its primary purpose was in diagnosing the needs of the individual student. The policy called for more diverse forms of assessment than the traditional examination. It placed a strong emphasis on continuous assessment, largely because continuous assessment can be used 'developmentally', in that teachers can teach to the demonstrated needs of each student (*ibid*: 9-10). Continuous assessment would be administered within schools, by teachers, with some external moderation – 'for example [sic!] by professional support services, within the guidelines set by the provincial education departments' (*ibid*: 11). The only external assessment would be at the end of Grade 9.

The implementation of this policy has been delayed by multiple logistical problems. The 1998 document itself recognised that teachers needed 'improved expertise...in designing, developing and using appropriate assessment instruments' (*ibid*: 10) and that the recommended range of continuous assessment techniques was demanding (*ibid*: 12, para 21), as was the expected record-keeping and reporting (*ibid*: 12-13, paras 22-25). The design of the assessment policy meant that it could not be implemented without a massive programme of teacher training. Such a programme is inherently desirable, and may in itself help to improve the quality of education, but it delays the implementation of a systematic assessment exercise. Moreover, the policy also required that expected levels of performance (EPEs) be specified for each grade in each of the three learning programmes in the Foundation Phase of schooling (concluding in Grade 9): five learning programmes in the Intermediate Phase and eight in the Senior Phase making a total of forty-eight sets of EPE. As was to become clear, this was too complex a policy to be implemented, especially in those schools which most needed systematic assessment.

In some documents, government departments had written of a series of national tests to be introduced in order to complement the matric examination. These would be held at Grades 3,6 and 9 (Department of Education, 2000b: 4; 2000c: 19). Confusingly, the Department of Finance has reported that the Department of Education planned to introduce new tests at Grades 3, 5, 7 and 9 (Department of Finance, 2000: 35). A series of national tests would obviously make it easier to monitor the performance of schools, teachers and students, to assess the effects of specific experimental reforms, and to take corrective action. In practice, however, the only national tests with systematic external moderation would be the Grade 9 'test'. Even this, according to the Chief Director for Quality Assurance, is to comprise 25% external examination and 75% continuous assessment. The national Department expects that the Grade 9 examination will be introduced at the end of the 2002 school year.[13] (Tests would be introduced at Grades 3, 6 and 9 levels as part of the Systemic Evaluation exercise - see below; these would be held for purposes of evaluating the education system, not student certification).

The Assessment Policy might help teachers, parents and students to monitor the progress being made by individual students, but it is very unlikely to provide reliable data on the performance of teachers or schools or educational reform programmes. It is over-reliant on teachers for implementation, and lacks adequate external monitoring. Moreover, it is in all likelihood too complex for many teachers in many schools.

Taylor (1999: 202) reports that the studies conducted for the President's Education Initiative found that 'the assessment skills of many teachers are rudimentary – to say the least'.

> 'Thus, the teacher questions observed by our researchers never moved beyond simple recall; pupils are seldom encouraged to ask questions themselves; peer interactions which focus on conceptual knowledge are rare; individual reading is almost non-existent, and written work is sparse and hardly ever goes beyond single words. Our research results indicate that the more sophisticated forms of assessment are beyond the reach of the majority of South African teachers at this stage.'

The National Policy on Whole School Evaluation

In June 2000 the Department of Education unveiled its new National Policy on Whole School Evaluation (Department of Education, 2000b). The new policy was proclaimed to be 'radically different from the previous school inspections' of the apartheid era. The focus would henceforth

[13] Dr Mgijima, interview, Pretoria, 13 July 2001.

be on the 'whole school', not individual teachers; the policy would be developmental, not punitive; and there was a commitment to transparency and consistency.

The background to the policy is the collapse of the discredited system of school inspections. In the early 1990s defiant school teachers, mostly in schools with black students, brought school inspections to an end. During the rest of the 1990s there was almost no system for evaluating schools or teachers. When the national Department of Education conducted a 'Quality Assurance Audit' in 1998 and 1999, with funding from the British Department for International Development, it found that the Professional Support Services (i.e. the former Schools Inspectorate) was dysfunctional at the provincial level. It was focused entirely on examinations and bureaucratic tasks, rather than developmental ones, although its role was supposedly now to supervise in order to enhance the quality and efficiency of basic education (see Department of Education, 1999).

The new policy on 'whole school evaluation' was developed against the backdrop of tortuous negotiations over the reintroduction of a system of teacher appraisal. The national Department of Education had sought to reintroduce teacher appraisal through a Developmental Appraisal System, in which appraisal was divorced from the processes of salary determination. In other words, appraisal was intended to identify teachers' strengths and weaknesses although it would have no consequences for individual teachers' salaries. This system of appraisal clearly lacks teeth. However, the national Department suggests that the new system allowed the gentle reintroduction of school visits and has paved the way for a future system that might have more teeth. Negotiations are now underway in the Education Labour Relations Council (ELRC) between the departments and the teachers' unions over how to link promotions to performance.[14] In practice, however, there are parts of the country where teachers do not allow even their school principal into their classrooms to observe their performances. Officially, the unions supposedly say that any such incidents should be reported and that a joint department/union team should investigate and take whatever action is necessary. The reality seems to be that there is still little or no observation, yet alone assessment of, teachers' performance in some parts of the country.

The proposed 'whole school evaluations' would be the centrepiece of a new system of developmental quality assurance. To secure union agreement, however, the policy has been formulated so that the evaluations lack bite. The policy would combine external evaluation and support with self-evaluation by school staff and governing bodies. Schools are supposed to produce regularly their own written reports for publication on a regular basis and are to be visited intermittently by external review teams. The policy does *not* entail any external assessment of learner achievement. The external 'supervisors' are required to observe lessons during their visits, but their visits have to be scheduled through prior agreement with the schools, allowing ample opportunity for teachers and schools to put their best face forward. Much of the evidence on what schools are actually doing comes from what schools say they are doing. The development of performance indicators is central to the policy, and the published 'Evaluation Guidelines and Criteria' provide useful guidance on how schools should be going about their work. Self-evaluation against these criteria as well as external observation and interviews might reveal whether schools are being managed and are operating in appropriate ways. These external evaluations might play a valuable developmental role. But the absence of any external evaluation of learner achievement and the prohibition on unannounced external visits sets tight limits on their utility. Schools may *appear* to be doing the right things while at the same time as they providing shockingly low-quality teaching.

[14] Duncan Hindle, interview, Pretoria, 13 July 2001.

Doubts have been raised over the affordability of the policy. If schools are to be evaluated every five years, then some five thousand or so schools need to be evaluated per year. If each evaluation takes a team of 'supervisors' several weeks to conduct (including preparation and the writing up of research), then the implications for personnel needs are daunting. At an estimate of, say, four weeks per school, and five supervisors per team, and taking into account the constraints of term time, a total of 2 500 specialist supervisors will be required (not counting administrative support or travel and living expenses). These external reviews are to be funded out of conditional grants from the national Department (which will also maintain a computer database containing all relevant information). The cost might be small compared with the total cost of the schooling system, but it would absorb most of the conditional grants currently budgeted under the government's Medium Term Expenditure Framework (MTEF).

By mid-2001 the provinces (excepting the Western Cape) were reported to have appointed between 250 and 300 people for the review teams, and trained them on the Guidelines and Criteria. The instruments to be used have been piloted. The Western Cape has opposed the system, and is preparing its own policy.

Systemic Evaluation

Some external evaluation of learner achievement is provided for in the third strand of quality assurance: systemic evaluation. The need for systemic evaluation was identified in the 1998 Assessment Policy (Department of Education, 1998). After some delays, the national Department produced in early 2001 successive drafts of a Framework for Systemic Evaluation, and held a colloquium in April of the same year. The Systemic Evaluation will involve testing of samples of students in Grades 3, 6 and 9 to see whether they are achieving the targets set in the Assessment Policy. The purpose is to measure 'the effectiveness of the entire educational system' (Department of Education, 2001a). As Dr Mgijima (the Chief Director for Quality Assurance in the national department) has put it, Systemic Evaluation seeks to answer the question, 'How well and to what extent do learners master basic skills, especially in the areas of literacy, numeracy and life skills and what are the factors influencing the acquisition of these skills?' (Department of Education, 2001b: 12).

The Minister of Education described 'Systemic Evaluation' of the education system as a whole as a 'national imperative'. Speaking in early 2001, he noted that South Africa was making 'huge investments in the education sector'. The government was increasingly concerned, he said, with the quality of education, and hence the evaluation of performance. He went on to say that the government, 'needs access to accurate information on the operations of the education system to determine whether the state is getting good value for its investment…The need for collecting data, evaluating the efficiency of our programmes and fostering broad debate on these bases to guide policy and decision-making, has thus become even more acute than before' (Department of Education, 2001d: 6-8).

The policy suggests that the initial, introductory rounds of the evaluation would involve a sample of 5% of pupils. A sample of grade 3 students would be tested in 2001, of Grade 6 students in 2003 and of Grade 9 students in 2005. Students from a range of schools must be included, but no more than forty pupils from any one school should be included in the sample. Thereafter, the policy implies (Department of Education, 2001a: 15), every student in the relevant grade would be tested.[15] Grade 3, 6 and 9 students would be tested every three or five years (the policy framework is inconsistent on the interval: contrast pages15 and 18). Students, teachers and parents would complete questionnaires, to provide a range of contextual data.

[15] It has been suggested that it was the Minister's intention to conduct a census-like test of all students in the particular grades in the selected years, but the logistical problems were daunting.

It is not clear that this Systemic Evaluation has been fully thought through. Testing a sample of students in different grades at regular intervals is more than adequate if the sole purpose is to assess the performance of the education system as a whole, as the document claims. Indeed, if sufficient additional or supplementary information were to be collected by means of questionnaires or through EMIS, there would be considerable scope for unprecedented analysis of the schooling system. The choice of interval between rounds of evaluation should presumably be made in such a way as to allow evaluations to track a cohort of pupils through the school system.[16] Why then consider testing *all* students? This only makes sense if the test results are to be used to monitor the achievement of individual students, i.e. they are performing the function of an examination, whether for public certification or to diagnose individual students' needs for remedial action.

The Department also proposes to have the tests administered by teachers from the schools concerned, with marking and coding done at district level, using 'identified moderators'. This clearly reduces costs, and might help to direct teachers (and district officials) towards the desired goal of improving the performance of their students. But there is a danger that the integrity of the testing process would be compromised, especially if the results are to be accessible to officials or the public, and there must also be doubts about the capacity of many schools and district officials to complete the process satisfactorily.

A Grade 3 test instrument had already been piloted in late 2000, and was due to be implemented in September 2001. Implementation might already have begun had the Department of Education decided to use test instruments already developed for Grades 3 and 6 by JET and other NGOs. The Department opted instead to commission the HSRC to develop a new set of instruments; it is reported that the JET instruments were not considered to be appropriate for South Africa's Outcomes-Based Education system. What the differences are between the test instruments is unclear. The contextual questionnaires are being developed by RIEP. The plan is that Grades 6 and 9 test instruments will be piloted in 2002 and 2004 respectively.

Test-Based Indicators

Lessons should be learnt from the assessment of pupil achievement through tests conducted as part of studies such as the Third International Mathematics and Science Study (TIMSS) series, and the Monitoring Learning Achievement (MLA) project. The experience of these shows that there are presently huge variations in the numeracy and literacy levels of students in the same grade, whilst overall test scores, especially in numeracy, maths and science, are shockingly poor. Indeed, many South African students fail to get onto the range for which the international test instruments were designed! Pupils in Northern Provinces, for example, also performed at far lower levels than students in Gauteng or the Western Cape.

Since 1993 a series of studies has collected data on the capability of representative samples of students. The first of these studies was the Project for Statistics on Living Standards and Development (PSLSD) survey, conducted by the University of Cape Town together with the World Bank, in the second half of 1993. One part of the PSLSD survey involved a short test of reading comprehension and numeracy. The PSLDS data has been analysed fully elsewhere (Fuller *et al*, 1995; Moll, 1998). In summary, it showed that 'the average African secondary school-leaver, with

[16] Thus if Grade 3 students are tested in (say) 2001, then Grade 6 students should be tested in 2004. The next round of Grade 3 tests should not be conducted in 2004, because that would involve testing them and Grade 6 students in the same year, causing problems for data capture and analysis. So the next round of Grade 3 tests should presumably be conducted two or four years after the first, in 2003 or 2005.

twelve years' schooling, only narrowly passes a primary-level literacy test, and primary school-leavers fail it miserably,' (Moll, 1998: 263).

Since 1993, more detailed tests have been conducted with much larger samples of pupils. These include TIMSS and TIMSS-R in 1995 and 1998 respectively, the Longitudinal Survey of Scholastic Achievement begun in 1995, and the MLA survey in 1999. The published results of TIMSS and TIMSS-R focus on pupils in Standards 5 and 6 (Grades 7 and 8). The Longitudinal Survey examines Grade 9 (Standard 7) pupils. The MLA survey focused on pupils in Standard 2 (Grade 4).

The Third International Mathematics and Science Study (TIMSS) was conducted in 1995 in 41 countries. TIMSS examined the science and mathematics curricula and measured student ability in these subjects, with a total of more than 500 000 students being tested. South Africa was the only participating country in Africa. Internationally, TIMSS focused on the following three student populations:

- population 1: the two adjacent standards containing the largest proportion of students aged nine years;
- population 2: the two adjacent standards containing the largest proportion of students aged thirteen years;
- population 3: the final standard of secondary school.

In South Africa, the Population 1 survey group was omitted ('owing to the underlying problems with language and the medium of instruction', as well as financial constraints - Howie and Hughes, 1998: 3-4); testing was thus limited to students in Grades 7 and 8 (Standards 5 and 6, i.e. the middle school years) and Grade 12 (Standard 10, i.e. the final year of school). Almost 3 000 Grade 12 students and nearly 10 000 Grade 7 and 8 students were tested, in a randomly selected national sample of schools.[17]

In 1998 South Africa participated in a repeat of TIMSS, known as the Third International Mathematics and Science Study-Repeat (or TIMSS-R). The other 37 countries that participated included a number of middle-income countries in East and South-East Asia, Latin America, Eastern Europe and the Mediterranean. South Africa was the only participating country in sub-Saharan Africa. The South African research was conducted by the HSRC. Unlike the original TIMSS, TIMSS-R was conducted only among Standard 6 (i.e. Grade 8) pupils. Just over 8 000 pupils in 194 schools were tested.

A Longitudinal Survey of Scholastic Achievement was reportedly conducted among 25 000 Grade 9 pupils in 1995 and again in 1996 (and was due to be repeated in 1998). Tests covered mathematics, science and English (Hartley *et al*, 1998a: 4). I have not obtained a full report on the survey or the results. Hartley *et al* report the 1995 results only by province. This allows for a once-off inter-provincial comparison, but not for the longitudinal analysis promised by the project title.

In August 1999, the South African Monitoring Learning Achievement (MLA) Survey tested Grade 4 pupils in literacy, numeracy and life skills. The study was commissioned by the national Department of Education as part of the Joint International UNESCO/UNICEF monitoring learning-achievement project. About 11 000 students in total were tested at about 400 schools (as well as via structured questionnaires to parents, principals, teachers and pupils) (Department of Education, 2000a: 14).

[17] There is some confusion over the precise numbers. The figure of 15 000 has been given for students being tested. Almost 1000 Grade 12 students were not included in the published analysis because they were repeating Grade 12 (Howie and Hughes, 1998: 21).

The study found that pupils' performance was 'poor'. The average score in each of the numeracy, literacy and life skills tests was below 50%. Table 4 shows the test scores for three individual provinces in and South Africa as a whole. There were big differences between provinces. The average score in the Western Cape, for example, was 38% compared with a low 26% in the Northern Province.

Table 4: MLA Test Scores, 1999

		Average score	**% with score 0-25**	**% with score 25-50**	**% with score 50-75**	**% with score 75-100**
Numeracy	SA	30	43.9	45.8	8.8	1.5
	Western Cape	38	28.3	47.7	20.1	3.9
	Eastern Cape	31	38.9	48.6	10.9	1.7
	Northern Prov.	26	54.8	39.9	4.7	0.7
Literacy	SA	48	13.3	47.1	26.8	12.8
	Western Cape	61	6.4	28.1	36.6	28.9
	Eastern Cape	48	9.4	48.3	32.6	9.7
	Northern Prov.	43	17.6	53.8	23.7	4.9
Life skills	SA	47	8.3	53.8	31.7	6.0
	Western Cape	56	3.7	36.2	48.8	11.3
	Eastern Cape	48	8.6	47.8	37.4	5.7
	Northern Prov.	45	9.1	67.7	20.5	2.7

Source: Dept of Education (2000a): Tables 18-20 and Figure 10

The MLA study, like TIMSS and other studies, supplemented the test-score data with other data on the students, including their school and home backgrounds. This should allow for a much more detailed analysis than hitherto of the factors contributing to educational success or stagnation. Among the factors already identified are high repetition rates (17% of pupils had repeated one or more grades) which in turn contributed to a variety of ages in one class.

Perhaps more important than the mean performance of students is the distribution of educational results. Is there equity in South African education in terms of performance? One particular element in public discussions of the matriculation results is that, in a small number of schools, they are appalling. Even in 2000, amidst greatly improved results overall, some fifty-six schools failed to register a single matric pass. Almost five hundred schools had a pass rate below 20% (although the number of schools with pass rates between 0 and 20% halved, from 940 to 499, between 1999 and 2000).

The MLA and TIMSS/TIMSS-R data allows for a more detailed analysis of the distribution of educational achievement among South African students. Table 5 above shows that in South Africa as a whole, 44% of the Grade 4 pupils tested got scores below 25%, whilst a small proportion got over 75%. In the Western Cape, the proportion of poor performers was much lower, and of high performers much higher, but there remained a wide range of ability. The TIMSS data provides a similar picture. Table 5 below indicates the range of test scores in three selected provinces. The difference between lowest and highest scores was enormous in every province.

Table 5: Range of Results in TIMSS, 1995

		n	average score	lowest score	highest score	standard deviation	% below 250 pts
St 5 maths	Western Cape	301	358	110	577	57	2.3
	Eastern Cape	515	341	179	577	58	5.6
	Northern Province	1060	328	110	523	51	5.7
	SA						4.5
Std 6 maths	Western Cape	84*	408	284	583	82	0
	Eastern Cape	546	341	196	583	56	4.4
	Northern Province	803	383	183	542	50	3.6
	SA						3.8
Std 5 science	Western Cape	301	349	128	587	57	11.6
	Eastern Cape	515	304	50	579	81	23.5
	Northern Province	1060	286	50	595	76	30.0
	SA						23.4
Std 6 science	Western Cape	84*	428	206	664	99	3.6
	Eastern Cape	546	307	94	664	79	20.7
	Northern Province	803	281	60	664	79	34.6
	SA						23.3

Note: * indicates a small n
Source: Hartley *et al,* 1998a, 1998b, 1998c.

Data on the distribution of results should also be available from the Longitudinal Survey of Scholastic Achievement, but this is not reported in Hartley *et al* (1998a).

In combination, these data-sets should make it possible to pinpoint what factors contribute to educational achievement. The PSLSD, TIMSS, MLA and Longitudinal Survey each collected data on pupils' school and home environments. Careful use of these data-sets should enable us to go beyond the preliminary studies undertaken by (for example) Case and Deaton (1999).

Conclusions

It is clear that matriculation results on their own cannot provide an adequate measure of the performance of schools or the utility of different educational programmes, and repetition rates are also an unreliable indicator. The national Department of Education recognises this, and has certainly been moving in the right direction with its suite of new Quality Assurance policies. There is a clear need for effective monitoring of individual students' progress, of the performance of schools as whole units, and of the education system as a whole. The Department's policies have surely helped to put a concern with performance, quality and even accountability to the fore of the educational system.

The suite of forms of assessment that are being developed by the national Department of Education will provide a range of valuable information. As discussed above, this suite comprises the following:

- the existing matriculation examination: this provides an indication of the performance of the one-third to one-half of South African children who reach Grade 12;

- the national Assessment Policy for students: this provides detailed information, through continuous assessment and a public examination (with some external moderation) at the end of Grade 9, in order to assess the performance and progress of individual children;
- 'systemic evaluation' of the schooling system: testing of samples of pupils in specific grades across the country, in order to assess the performance of the education system as a whole; and
- 'whose school evaluation' of individual schools: a combination of external and self-evaluation, but without external assessment of pupils' performance, in order to assess the performance of individual schools; this is an essentially qualitative evaluation.

However, there are clear shortcomings in the Quality Assurance system being put into place. There is no sign of any integration or co-ordination between Whole School Evaluation and Systemic Evaluation. How, for example, is the Grade 9 examination (in the student Assessment Policy) linked to the proposed Grade 9 tests (under the Systemic Evaluation policy)? (See Department of Education, 2001b: 31-2). The absence of any external testing of learner achievement in Whole School Evaluation is a serious flaw. Monitoring what schools are doing (or, more realistically, what they want supervisors to see or think they are doing) is no substitute for monitoring what schools are achieving in terms of the measurable progress of their students. The purpose of investment in education is to enhance literacy, numeracy and other essential skills. External monitoring of schools' progress in developing these skills must be a core component of any evaluation.

Co-ordination between Whole School Evaluation and Systemic Evaluation would help to overcome this problem. Testing student achievement might be done as part of the Systemic Evaluation process in selected schools at much the same time as those schools participate in Whole School Evaluation. The testing would have to be monitored carefully. This would allow the test results to be analysed together with the detailed and essentially qualitative results of the Whole School Evaluation process. The combination of the two processes would generate information which would be valuable in assessing performance, diagnosing strengths and weaknesses and indicating what interventions, support or initiatives would enhance performance in future. The 'macro' analysis provided by Systemic Evaluation would be linked to the 'micro' analysis of Whole School Evaluation.

In practice, the lack of integration between the different components of the new Quality Assurance system might reflect the limited staff and resources of the Chief Directorate for Quality Assurance in the national Department of Education, and possibly the provincial Quality Assurance units also. Conducting a thorough programme of quality assurance will require increased capacity – including the capacity to analyse data – at both national and provincial levels, both inside and outside government departments.

The national Department of Education is clearly concerned about the cost of implementing quality assurance. It is germane to point out that independent education programmes funded by business or foreign donors typically dedicate between 1 and 2% of their budgets to appraisal, and this is arguably insufficient. Nonetheless, 1% of the consolidated spending on primary and secondary schooling would amount to a very large sum indeed: R500 million or so, per annum.

A crucial question facing policy-makers and administrators is: which policies 'work' and which do not, in terms of building students' skills? How should resources be spent (or, more accurately, invested) if they are to achieve the best possible outcomes in terms of students' skills? How can we learn from the experience of spending R50 billion or so on schools each year? What can we do – and what should we be doing – so that we are confident that public resources are being put to the best possible use? The envisaged suite of quality assurance mechanisms would

generate, if they were conducted in an exemplary manner, much of the information required to identify what 'works' and what does not. The results of the Grades 9 and 12 examinations, together with the results of testing as part of systemic evaluation, would provide some indication of the performance of individual schools. The Whole School Evaluation would provide some information on what schools are actually doing, which might reveal the differences between schools that produce good results and those that produce poor ones (however they are measured, and taking into account family and community factors). And the information on individual students would provide a very detailed 'micro' picture of what this means for individual children.

However, there are problems with the suite of mechanisms. Systemic Evaluation takes place at too 'macro' a level; Whole School Evaluation is inherently flawed; the Assessment Policy is too 'micro'. More specifically, the contribution of this suite of forms of assessment will be undermined by two major problems. Firstly, there is insufficient regular, external testing assessment of pupils. Secondly, it will be no easy task to ensure that the information provided by Whole School Evaluation is both of high quality and usable. Both problems stem from the scale of the exercise: an estimated 2 500 specialist supervisors, across the country, visiting 27 000 schools at regular intervals and writing up their reports. The numbers of people involved are perhaps of less concern, since they will be playing a valuable developmental role, advising schools and teachers on how to do better. But the number of teams of supervisors visiting schools means that there will be a flood of reports to provincial and national departments, and it is far from clear what bureaucratic procedures and structures will be capable of making sense of this.

In short, whilst the suite of forms of assessment might provide some data on performance and some information required for decentralised developmental efforts (i.e. for teachers to assist students whose progress is slow and for supervisors to assist schools and teachers whose practice is deficient), it is unlikely to provide sufficient data on what 'works' and what does not. These forms of assessment are necessary but not sufficient. In assessing *how* to improve students' performance, *experimental* programmes and projects are indispensable. Such programmes involve concentrating assessments on small numbers of schools that have been selected as the sites of focused interventions or reforms. The Department of Education's suite of forms of quality assurance is one designed for an education system that is basically working well, and which simply needs some tinkering at the edges; it is not designed to deal with an education system that is in crisis in many respects in crisis, and which probably requires radical interventions and reforms.

Careful planning, implementation and evaluation in experimental programmes should provide lessons that save time and money when broader reforms are attempted. Sometimes educational programmes are not designed as experiments, but nonetheless may serve this purpose. Conversely, projects may be set up as experiments, although they have no value due to careless design or shoddy evaluation. It must be emphasised that experimental programmes are not a *substitute* for the kinds of assessment that the various departments of education have set up or are planning; experimental programmes cannot provide the important information that other forms of assessment generate. Rather, they are a complement, providing different information for different purposes.

References

Anderson, K., Lam, D. & K. Hillard. 2000. *Grade Repetition, Schooling Attainment, and Family Background in South Africa.* Presentation of preliminary results at seminar at the University of Cape Town, April.

-----, Case, A. & D. Lam. 2001. Causes and Consequences of Schooling Outcomes: Evidence From Survey Data. *Social Dynamics 27(1)*.

Baird, J-A. 2001. Comments on the Pilot. Paper presented at the *Colloquium on Systemic Evaluation*. Benoni, April.

Case, A. & A. Deaton. 1999. School Inputs and Educational Outcomes in South Africa. *Quarterly Journal of Economics*, *458,* 104710-84.

Crouch, L. 1998. *Education Data and the 1996 Census: Some Crucial Apparent Problems, Possible Strategies for Resolution*. Mimeo, December.

----- .1999. *Education Indicators in South Africa: Who Needs Them? What For? Some Theoretical and Practical Considerations*. Mimeo, April.

----- & T. Mabogoane. 1997. Aspects of Internal Efficiency Indicators in South African Schools: Analysis of Historical and Current Data. *Edusource Data News 19,* 4-28.

----- & ----- . 2001. No Magic Bullets, Just Tracer Bullets: The Role of Learning Resources, Social Advantage, and Education Management in Improving the Performance of South African Schools. *Social Dynamics*, *27(2)*.

Department of Education. 1998. Assessment Policy. Published in *Government Gazette,* 402 (19640), 23 December.

-----. 1999. *Quality Assurance Audit Phase 2*. Two volumes, co-ordinated, compiled and written by Khosa, G. & S. Motala, March.

-----. 2000a. *Education for All: The South African Assessment Report*. March.

-----. 2000b. *The National Policy on Whole School Evaluation* (June 2000). Downloaded from departmental website, 3 January 2000.

-----. 2000c. *Annual Report 1999*. Pretoria: Government Printer, RP 62/2000.

-----. 2000d. Report on the Senior Certificate Examination. 28 December.

-----. 2001a. *Education Statistics in South Africa at a Glance 1999.*

-----. 2001b. *Report on the Colloquium on Systemic Evaluation.* Draft Report. Lopanong Conference Centre, April.

Department of Finance. 2000. *Intergovernmental Fiscal Review 2000*. Pretoria: Department of Finance.

Financial and Fiscal Commission. 1998. Report for the United Nations.

Fuller, B., Pillay, P. & N. Sirur. 1995. *Literacy Trends in South Africa: Expanding Education While Reinforcing Unequal Achievement.* University of Cape Town: South African Labour and Development Research Unit.

Gilmour, D. 1999. School Performance Indicators in Developing Countries. Unpublished paper.

Hanushek, E., *et al.* 1994. *Making Schools Work: Improving Performance and Controlling Costs*. Washington DC: The Brookings Institute.

Hartley, A., Visser, M. & C. Sheppard. 1998a. *The Condition of Primary and Secondary Education in the Western Cape (1995/1996)*. Pretoria: Human Sciences Research Council.

-----, ----- & ----- 1998b. *The Condition of Primary and Secondary Education in the Eastern Cape (1995/1996)*. Pretoria: Human Sciences Research Council.

-----, ----- and -----. 1998c. *The Condition of Primary and Secondary Education in the Northern Province (1995/1996).* Pretoria: Human Sciences Research Council.

Howie, S. & C. Hughes. 1998. *Mathematics and Science Literacy of Final-Year School Students in South Africa: A Report on the Performance of South African Students in the Third International Mathematics and Science Study (TIMSS)*. Pretoria: Human Sciences Research Council.

Lam, D. 1999. Generating Extreme Inequality: Schooling, Earnings and Intergenerational Transmission of Human Capital in South Africa and Brazil. *Research Report No. 99-439*. University of Michigan: Population Studies Centre at the Institute for Social Research.

Moll, P. 1998. Primary Schooling, Cognitive Skills and Wages in South Africa. *Economica, 65,* 263-284.

Muller, J. 1999. Reason, Reality and Public Trust: The Case for Educational Research for Policy. In Taylor, N. & P. Vinjevold (eds) *Getting Learning Right*. Johannesburg: Joint Education Trust.

Porteus, K. 2000. The Personal and the Global: Grinding Gears? Reconsidering Efficiency Debates in General Education. *Quarterly Review of Education and Training in South Africa, 7(3)*.

South African Institute of Race Relations. Various years. *South Africa Survey*.

Tan, J., Lane, J. & G. Lassibille. 1999. Student Outcomes in Philippine Elementary Schools: An Evaluation of Four Experiments. *World Bank Economic Review, 13(3),* 493-508.

Taylor, N. 1999. Learning Outcomes: How Will We Know Them? In Taylor, N. & P. Vinjevold (eds) *Getting Learning Right: Report of the President's Education Initiative Research Project.* Johannesburg: Joint Education Trust.

----- & P. Vinjevold. 1999. (Eds) *Getting Learning Right: Report of the President's Education Initiative Research Project*. Johannesburg: Joint Education Trust.