Accountability and Support in School Development in South Africa

Nick Taylor

JET Education Services

Paper presented to the 4th Sub-regional Conference on Assessment in Eduction

26-30 June 2006

Hosted by UMALUSI

1. Introduction

Around the world schools are under two kinds of social demands: the press for increased access, and the demand that graduates with ever more sophisticated cognitive skills. These pressures are frequently in conflict with one another. South Africa is no exception and, while we have done relatively well in providing school places to close to 100% of children of primary school age and an increasing proportion of children of high school age, there is also evidence that this press for equity has come at the expense of quality (Crouch and Fasih, 2006). As the South African economy begins to grow at levels last seen decades ago, it is clear that a lack of skills is exerting a brake on both national development and the social mobility of poor South Africans. The quality of schooling is key to the quality of skills in every sphere of social and economic life. This paper examines recent efforts in South Africa to improve school quality through the twin tools of accountability and support.

2. Standards-based Accountability

Standards-based accountability[1] (SBA) has assumed increasing importance in many countries over the last two decades or more, as a means of improving school outcomes. According to Carnoy et al (2003), by 2001 49 states in the US had adopted some form or other of this approach. The assumptions underlying standards-based accountability are:

- clearly defined standards, in the form of a common curriculum, explicate to teachers and pupils what is to be learnt;
- state-wide or national tests assess the extent to which schools and pupils are achieving the standards;
- rewards and sanctions accompany the results of the tests: for pupils these occur in the form of access to further and higher education and the job market, thus having an important bearing on career choices; for schools the

[1] Also known as performance-based reform (Hopkins, 2002) or outside-in reform (Muller, 2000).

> consequences come in the form of market information, signaling the quality of the school to parents, or via direct rewards and sanctions administered by the district, state or national bureaucracy.

In the face of its popularity with education administrators, SBA has generated strong criticism from academic circles, where four arguments against this approach are raised. First, it is argued, the high-stakes nature of the testing component leads to distortions in the curriculum. The main kind of distortion arises from the form which large-scale tests sometimes adopt (multiple-choice, solving problems of a restricted nature, individual responses), leading to a focus on certain parts of the curriculum and a neglect of others (extended writing, solving complex problems, teamwork). This narrowing of the curriculum is further exacerbated when testing is restricted to certain subjects or certain grade levels of the school, and when teachers 'teach to the test', drilling pupils on the kind of items commonly encountered in the tests, at the expense of other parts of the curriculum.

A second common criticism of SBA is that it takes no account of the value which a school adds. This line of argument arises from the strong correlation between educational outcomes and socio-economic status (SES). For example, a school situated in an affluent area and attended by pupils from middle class homes may be underperforming relative to its SES peers, but easily outscores a poorer school which is doing a really good job with children from working class homes. The latter school would clearly be at a disadvantage if judged on a common 'league table' in which no account is taken of SES, while the underperforming high-SES school would look good.

A third criticism leveled against the standards-based approach is that it generates high levels of anxiety among children, for whom success or failure in one set of tests at one point in their lives will have a profound influence on the rest of their careers.

Carnoy et al (2003a; Elmore, 2003) raise a fourth problem with accountability systems linked to high-stakes tests, which is perhaps more fundamental than the previous three lines of argument. These authors conclude that such systems are effective in raising aggregate student scores, but that there is wide variability among

schools in their specific responses to standards-based reform initiatives. They argue that the response of any particular school to *external accountability measures* depends on the state of its *internal accountability systems.* The latter are defined as the collective expectations held by members of the school community, together with the organizational rules, incentives and implementation mechanisms that constitute the school's formal accountability system (including instructional practices and supervision processes).

Strong internal accountability systems enable a school to respond positively to standards-based reform programmes, but schools whose internal systems are weak are unable to respond. The majority of schools, according to Carnoy et al (op cit), fall into the latter category, and the key to improving their ability to respond is capacity building, aimed at aligning and strengthening internal accountability systems.

These ideas frame our examination of the evidence for the effects of SBA and capacity building (support), in various combinations, on school performance in South Africa over the last decade.

3. SBA can have significant positive effects on targeted indicators

A feature of the post-1994 profile of SC results was a steady decline in both the pass rate and the proportion of pupils attaining university exemption. While the number of candidates fluctuated between 450 000 and 550 000 over the period 1994 to 1999, the pass rate declined from the 58% to 49%, and the exemption rate dropped from 18% to 12% (see Table 1). This decline should not be surprising, given the thorough-going reorganisation of the entire school system following the change in government in 1994 and the consequent destabilization this must have had on schools. However, after the second general election of 1999 government began to pay serious attention to the question of Senior Certificate (SC) results, the only quality indicators available, then and now. Among a number of measures directed at improving the functionality of schools and the implementation of the new curriculum and other policies, the national DoE established a National Monitoring Forum, the aim of which was to co-

ordinate improvement in the SC exam results (MoE, 2000). Each province was required to institute a SC improvement plan with a special focus on underperforming schools, defined as those which achieved pass rates in the 0-20% category. The results of these efforts was immediate, with pass and exemption rates showing a dramatic turnaround in 2000 and a sharp upward trajectory in subsequent years.

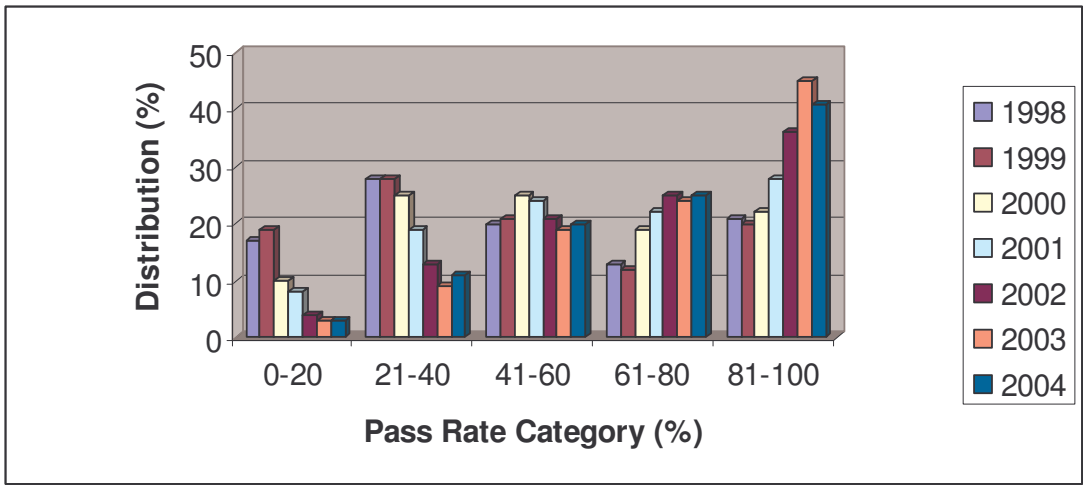
Table 1: Senior certificate examination results, 1994-2002

	Candidates	Total Passes	%	University Exemption	%
1994	495 408	287 343	58	88 497	18
1995	531 453	283 742	53	78 821	15
1996	518 032	278 958	54	79 768	15
1997	555 267	261 400	47	69 007	12
1998	552 384	272 488	49	69 856	13
1999	511 159	249 831	49	63 725	12
2000	489 941	283 294	58	68 626	14
2001	449 371	277 206	62	67 707	15
2002	471 309	324 752	69	75 048	16
2003	440 267	322 492	73	82 010	19
2004	467 985	330 717	71	85 117	18
2005	508 363	347 184	68	86 531	17

Source: DoE, 2004, 2005a, 2005b, 2005c

As we shall show later, the effects of these SBA measures were not felt equally in all schools. Nevertheless, they did effect schools in all pass rate categories, with the numbers of schools achieving 0-20% and 21-40% dropping sharply, and the numbers in the two highest deciles increasing, and the middle decile remaining more or less constant (see Figure 1).

Figure 1: Percentage distribution of schools by pass rate category, 1998-2004



Source: Collated from DoE, 2004, 2005a, 2005b.

The most intense SBA initiative at the provincial level was the Education Action Zone (EAZ) programme adopted by the Gauteng Department of Education (GDE) in 2000. Seventy schools in the province which exhibited pass rates below 20% were targeted for a package of interventions. The EAZ was designed as a comprehensive systemic initiative which would include monitoring schools and providing support and training to principals, teachers and pupils. However, in reality, the programme did not fully meet its systemic intentions, focusing largely on accountability measures (Fleisch, 2001; 2003; 2006). A project approach was adopted in administering the programme, rather than strengthening the systems and capacity for school monitoring and support in the standard line functions of the GDE: thus, the EAZ was managed from the provincial head office, with special units responsible for earmarked schools and reporting directly to the provincial Minister of Education and the Head of Department.

The EAZ achieved an impressive rise in SC results in targeted schools on a range of indicators: numbers of candidates passing at both higher and standard grades (HG, SG), overall pass rate, university exemption rate, and the numbers of A symbols achieved by pupils (80% or more on aggregate across all subjects) (Table 2).

Table 2: SC Results in Education Action Zone Schools, 1996-2003

	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Pass Rate (%)	24.32	17.16	20.49	15.59	33.09	48.59	59.19	66.48
University Pass Rate (%)	2.25	1.76	1.62	0.99	2.10	3.53	4.59	6.52
Number Passed Higher Grade	211	167	155	104	170	197	246	387
Number Passed Standard Grade	2015	1455	1467	1530	2508	2515	2024	3631
Total Passed	2226	1622	1622	1634	2678	2712	3170	4018
Number A Symbols Awarded	8	3	28	32	200	124	216	373
Number Wrote Higher Grade	8455	7921	6254	4252	1778	1526	1383	1571
Number Wrote Standard Grade	942	1551	3334	6229	6315	4055	3973	4468
Total Wrote	9397	9472	9588	10481	8093	5581	5356	6039

Source: Fleisch, forthcoming.

Not only are these results very impressive on their own, but EAZ schools also increased relative to other schools in the province: thus, in the first two years the aggregate pass rate for project schools increased by an average of 14,5%, which exceeded the improvements shown by both other former DET schools in the province (up 10,1%), and all public schools in Gauteng (5,3%) (Fleisch, 2001).

The evidence presented this far clearly shows that the pressure created by SBA measures can result in dramatic increases in targeted indicators on a large scale. The fact that the improvements shown by EAZ schools exceeded that of comparable sub-categories of schools in the province, suggests that the greater the pressure exerted by SBA, the greater the gains. Whether these indicators measure anything worthwhile, and whether increases in the indicators reflect improved performance are questions to which we now turn.

4. SBA may lead to curricular and other distortions

Skepticism has been expressed about the meaning of the changes exhibited both by the national trends shown in Table 1 and the EAZ results in Table 2. Indeed, in the very first year of the turnaround, Professor Jansen expressed doubt as to whether they reflected anything more than 'imaginative statistics' (Fast Facts, 2001, 3). Similarly, while conceding that improvements in the results of EAZ schools were spectacular, Fleisch (forthcoming) argues that, instead of reflecting real increases in quality, these changes may have been achieved through manipulations of a more cynical nature. Fleisch advances four kinds of alternative explanations: reducing the number of SC candidates, through the exclusion of high-risk pupils; encouraging pupils to write papers at the easier standard grade level; lowering the standard of examination papers; or changes in the moderation process. We examine each of these alternative explanations in turn.

4.1 Excluding high-risk candidates

There is no doubt that the improvements in SC results over the period 2000-2003 were accompanied by significant reductions in the numbers of pupils registered for

the exam in EAZ schools. The number of candidates dropped by 4 442 (42%) between 1999 and 2003. However, in the face of a concomitant rise of 2 384 (146%) in the absolute number of pupils passing, it seems obvious that clearing grade 12 classes of pupils who were unlikely to pass, allowed greater attention to be focused on more promising candidates, thus achieving not only very significant quantitative gains in educational opportunity for these pupils, but also vastly improved rates of efficiency.

4.2 Increasing the ratio of SG:HG passes

In querying the value of the improvements shown by EAZ schools, the second kind of argument raised by Fleisch is that the increases in pass rate may have been achieved by shifting the pattern of registration away from the HG level towards SG, thus making it easier to pass. As Table 3 shows, the data does not support this hypothesis.

Table 3: Pass rates compared with SG:HG ratios, EAZ schools, 1996-2003

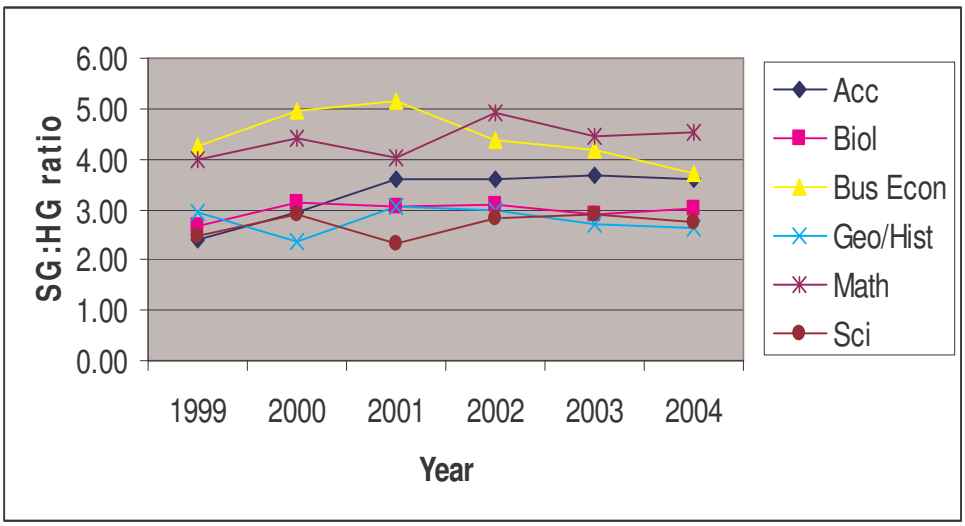
	1996	1997	1998	1999	2000	2001	2002	2003
Pass Rate (%)	24.32	17.16	20.49	15.59	33.09	48.59	59.19	66.48
SG:HG	9.55	8.71	9.46	14.71	14.75	12.77	8.23	9.38

Source: Calculated from Fleisch, forthcoming.

In the EAZ sample in Gauteng, a sharp rise in the SG:HG ratio from 1998 to 1999 was accompanied by a fall in pass rate; conversely, over the period 2000-2003 a steady rise in the pass rate was accompanied by falling SG:HG ratios.

Nor is there a clear association between increasing pass rates and rising SG:HG ratios in the selected subjects for which the DoE publishes national figures, as shown in Figure 2.

Figure 2: SG:HG ratios for selected subjects, national means, 1999-2004



Source: Calculated from DoE, 2004, 2005a, 2005b. Figures for Geography only published for 1999 and 2000; thereafter History is shown instead.

Against steadily rising pass rates, the SG:HG ratio for 5 of the 6 subjects did rise from 1999 to 2000, but after this only Accountancy continued to rise, while the ratios for three other subjects fell in 2001 while a fourth declined a year later. After this the patterns tend to flatten or decline slowly. The available evidence indicates that there is no systematic relationship between SG:HG ratios and rising pass rates.

4.3 Lowering the standard of the exam papers

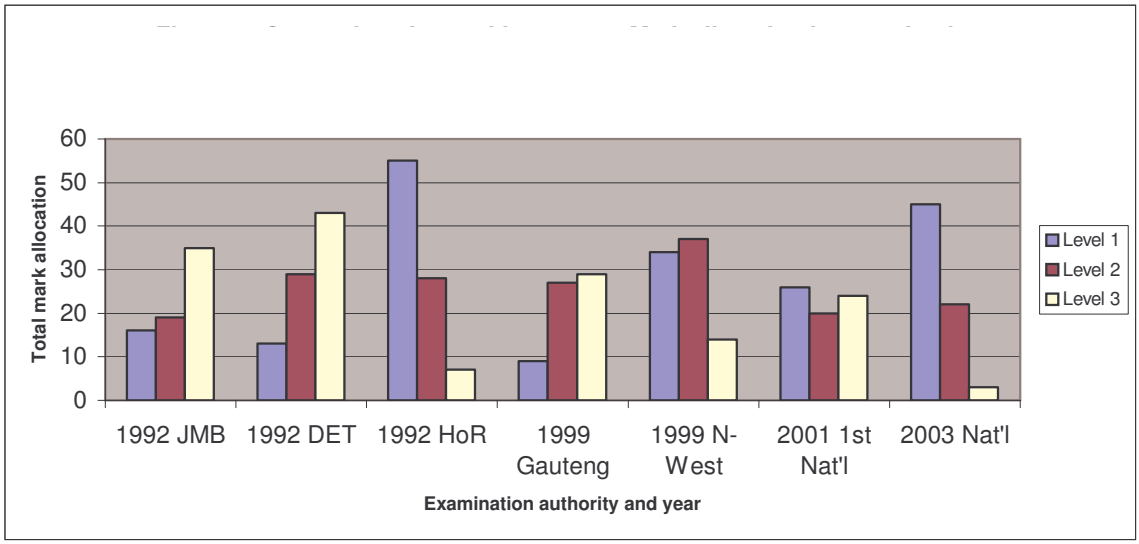
In response to growing public concern that the pass rate improvements since 1999 were the result of lowered standards, Umalusi commissioned a research project to investigate the issue. Question papers from 1992 to 2003 in Mathematics, Physical Science, Biology, History, English Second / Additional Language and English First Language were evaluated. A three-member team of experts was appointed for each of the six subjects to determine whether a discernable drop in standards had occurred over this period.

One of the specific questions posed to subject teams was whether the proportion of items representing different levels of cognitive challenge had changed. For the

purposes of this exercise, individual items were rated on the scale of 1 – 3, with (1) indicating the simplest and most basic question types and/or knowledge required to answer them, (2) indicating questions of average difficulty, and (3) questions involving more sophisticated linguistic, literary and general knowledge.

A finding which was common to the report of most of the teams was that the standard varied across the different papers set by various examining authorities over the period in question. This is well illustrated in the graph provided by the English Second/Additional Language team (ESL).

Figure 3: Comprehension and Language: mark allocation by authority and year.



Source: Yeld et al, 2004

A second point made very strikingly by the data shown in Figure 3 is the change in the proportions of item types between the 2001 and 2003 national ESL Language papers, with a rise in the number of easier questions (Level 1) and a fall in the number of questions requiring higher levels of understanding (Level 3). In conclusion, the team noted that 'From the analysis of the papers above, it is difficult to escape the conclusion that the nationally set paper (Paper 1) is becoming easier – or, in the jargon of the examiners, becoming 'more accessible' ' (Yeld et al, 2004, 14).

'Accessibility' had clearly become the watchword for examiners in a number of subjects, and the conclusion reached by the ESL team was confirmed, in one way or another, by all but one of the other subject teams. The Biology team noted a trend

toward setting a larger number of questions, each containing fewer marks, and that a consequence of this practice is that '… learners are seldom asked to demonstrate in-depth knowledge of important biological processes…', that 'The many short questions permit a fairly superficial coverage of all topics, but do not probe learners' ability to articulate their answers in a written form, nor to illustrate their answers appropriately', and that '…opportunities to test ability to synthesize information from different topics are limited' (Dempster et al, 2004, 14).

Although they had access to a wide variety of papers from different authorities, because of the limited number of papers which provided valid comparison points across time, the English First Language team decided to compare only the 1999 and 2003 papers set by the Gauteng Education Department. The team found an overall decline in cognitive challenge in the language paper, and to a lesser extent in literature. They noted that in 2003 very little reading or specific knowledge of English was required for language, while for literature many questions could be answered without knowledge of the full texts of a number of set works (Allais et al, 2004).

The History team noted with approval the approach adopted by the new curriculum, but concluded that, although assessors paid lip service to these principles, ' … it in fact did little more than reward rote learning under a different guise', and that ' … there is clearly less emphasis on critical thinking and more reliance on questions which identify tasks in terms of "describe" or "discuss" rather than in terms of "explain" or "analyze" ' (Kallaway et al, 2004, 10). As a consequence the team expressed strong feelings that ' … these phenomena had a considerable effect on the profile of marks received in 2003 and boosted the marks unjustifiably' (ibid, 12).

In mathematics the SG and HG papers for the 1999 and 2003 national examination were compared. At SG, for both papers 1 and 2, the team described a 'dramatic' decrease in the number of Level 3 questions in comparison to Level 1 items (Carrim et al, 2004). The same comparison was not done for the HG papers, although the report noted that in the national Paper 2 only 3% of marks were allocated to Level 3, compared with 25.5% in the corresponding IEB paper.

Science was the only subject in which the standard, as a whole, was found to be consistent over time: while the 2003 Physics paper was slightly easier than previous papers, the Chemistry paper was slightly more difficult (Masemula et al, 2004).

4.4 Changes in the moderating process

In addition to assessing the quality of exam papers, the Umalusi review also looked at the moderation of raw scores. The latter review noted that, because of the large differences in standard across pre-1994 examining authorities, many students were at a disadvantage in a common system. To compensate, adjustments performed at statistical moderation meetings from 1996 to 2001 were generally upward, frequently by the maximum of 10% allowed. Furthermore, ' … [o]ne of the consequences of the predominant trend of upward adjustments was a reluctance by the new examination authorities to accept downward adjustments when these were recommended by the SAFCERT / UMALUSI statistics team. … This certainly has resulted in an upward movement in pass rates' (Umalusi, 2004, 8).

4.5 Marking

One of the most important nodes in the assessment process is the scoring of scripts. In assessing the worth of student reponses, marking requires the exercise of high levels of professional judgement, in order to generate reliable results. And the higher the skill level required to answer any particular item, the higher the level of professional judgement required on the part of the assessor.

A second reason why marking is so important in any SBA system is because it provides opportunities for building capacity of teachers and communicating the standards to them. One of the best systems in this regard is the Key Stage testing process in England, where every child in the country is tested at key stages in her development. And a striking feature of the Key Stage system is the way in which teachers are trained to do the marking, and every marked script is sent back to the schools, with the marking memorandum, so that teachers and children in every

classroom in the country are clear on the criteria used to evaluate responses to every item in the tests.

South Africa has a long way to go before we reach these levels of organization and clarity of communication. But the question that concerns us here is: did manipulation during the marking process have a bearing on the increase in matric results in the period 1999-2003? There is insufficient evidence to answer this question, although the Umalusi research team for ESL examined 20 language scripts, drawn at random from one of the provinces, noted that the highest score was 30%, and wondered about the validity of the 98% pass rate in the subject across the province. This scrap of evidence raises serious questions concerning the training and quality control of markers, and the extent to which scoring is moderated and adjusted in districts and provinces, long before the marks get to the national adjudication process.

4.6 Conclusion

What are we to conclude from this evidence? Is it clear that the large rises in our indicators for the SC exam results were largely due to easier papers and mark manipulation? Certainly, the trend turned as soon as the Umalusi research was made public, and the standard-setting body wrested control of the moderation process from the DoE in 2004 and exerted closer oversight of the standards of the 2005 exam papers. These steps were accompanied by an immediate change in SC pass rates, showing a downturn in 2004 which was continued into 2005 (see Table 1), although the numbers of both ordinary and exemption passes continued rise.

But were all the gains of 1999-2003 due to the manipulation of standards and scores, or was there also some improvement in delivery by schools, either in tighter management or improved teaching, or both? Evidence for the effects of SBA on school level improvements is less conclusive than the strong case we have presented so far for the manipulation of results. What is blindingly clear is that the primary purpose of SBA is defeated if standards are not kept constant or improved over time. It is also crystal clear in the South African case that high school standards are very

low, focusing largely on cognitive skills of an elementary nature, at the expense of higher order processes: analysis, synthesis, and the development of written argument. The best intentions of any SBA system are to improve standards, but, if close watch is not kept on all the exits, then it can have the opposite effect, and our analysis so far indicates that the DoE succumbed to this temptation in the years 1999-2003.

The good news is that Umalusi stopped the rot, following its legislative brief to be the watchdog of standards. Whether it manages to improve the quality of education, in the face of all contrary temptations, will be a matter for history to judge. Umalusi's first task in improving the quality of schooling is to steadily improve the standards of the exam papers, accompanied by an explicit description of how this is to be done so that teacher and pupils are able to prepare for the new demands.

By far the most important step towards improving standards is to undertake a radical overhaul of the ESL curriculum. Over 80% of SC candidates both write this subject, and rely on the conceptual and language skills it provides to learn all their other subjects. Yet, the findings of the Umalusi research indicate that the current low standards of ESL are severely aggravating the educational disadvantages which ESL speakers already labour under because of poverty. Furthermore, the practice of compensating these candidates by adding 5% to all their non-language scores, instituted in 1998 as a temporary measure, is counterproductive to improving the quality of schooling. These are sensitive issues and, while taking the current easy route may bring short-term political gains, the long-term educational damage accretes year by year.

5. SBA and support together leverage significant improvements in school performance

The Quality Learning Project (QLP) was programme which provided support measures to schools, intended to complement the SBA effects of the SC regime in South Africa. Working in 524 high schools across the nine provinces provided training and support to teachers, principals and district officials. A longitudinal evaluation (HSRC, 2005) found that over the life of the project QLP schools achieved

significantly better results in the SC exam than selected control schools, in terms of greater numbers of overall passes, university exemptions, passes in mathematics, and in the pass rate. Similarly, improvements shown by QLP schools were significantly better than those of the national mean (Table 4).

Table 4: Comparison of QLP SC results with the national mean, 2000-04

	Increase 2000 – 2004								
	Passes		**Exemptions**		**HG maths**		**SG maths**		***% Pass Change***
	No	***%***	**No**	***%***	**No**	***%***	**No**	***%***	
Total QLP	4167	18.3	1182	34.8	585	152.3	8741	137.5	*14.0*
Total SA	47314	16.7	16493	24.0	8466	47.0	46512	58.0	*12.8*
Difference*		***1.6***		***10.8***		***105.0***		***79.0***	***1.2***

*Computed by subtracting the improvements exhibited by the national mean over the life of the project from those exhibited by QLP schools.
Source: Taylor and Prinsloo, 2005.

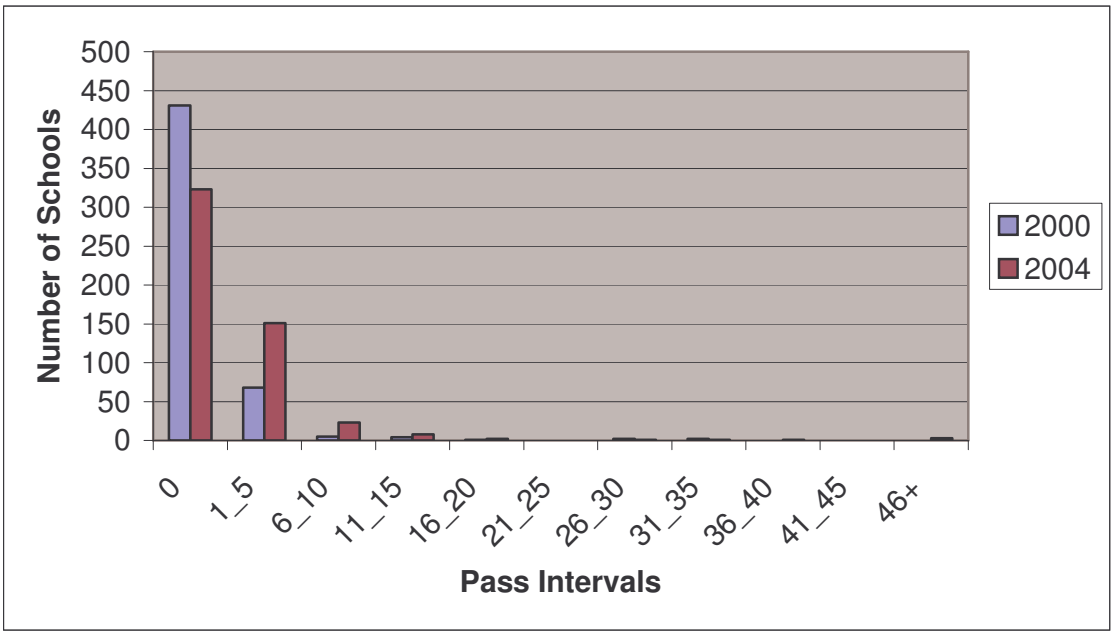
The evaluation concluded that the project had a significant effect on the performance of these very poor, largely rural schools: the QLP support programme resulted in improved practices in school management and classroom teaching, which in turn produced improved learning (HSRC, 2005). In Carnoy et al's (2003) terms, the support provided by the QLP assisted schools to strengthen their internal accountability systems, which in turn enabled them to meet the external accountability challenge posed by the intense SBA expectations of the time. Furthermore, we infer that the additional gains made by QLP schools, in comparison to the national mean, resulted from improved learning rather than any manipulation of the results, since the latter factor was constant for all schools.

Interestingly, the QLP evaluation also tracked pupil performance on math and language tests at grades 9 and 11. The only learning gains discernible were improvement in writing at grade 11. The most likely explanation for this disappointing result, in the light of the very impressive improvements at SC level, is that, whereas intense pressure is put on schools to perform in the SC exams, no such pressure is applied at lower levels of the system. Hence, capacity-building is most effective when accompanied by SBA measures.

6. SBA and support have no effect on some schools

Impressive as the QLP results shown in Table 4 are, they also have a dark side. This is shown in Figure 4, a frequency distribution of QLP schools, by number of HG maths passes achieved, before and after the intervention.

Figure 4: Comparison of HG maths passes in QLP schools, 2000 and 2004



These figures reveal that the improvement in maths results shown in Table 4 were achieved by only one-third of the schools. Fully two-thirds remained without a single HG maths pass at the end of the programme: these schools gained nothing, in terms of improving their HG maths passes, from the considerable resources expended on them over a period of 5 years.

These schools would be classified as 'failing' in the three-part classification devised by Hopkins et al (Hopkins, Harris and Jackson, 1997). Rewards and sanctions have no effect in these situations, as the schools are unable to help themselves: they require a high level of external intervention and support. There should be a clear and concerted focus on a specific, limited number of factors. In many schools in this state the first thing to be done is to remove the principal; often strong mediation is required to break

situations of conflict between various groups in the school. Only government has the authority to intervene here.

One of the most important lessons of the QLP is that programmes of this kind are only successful in schools which have a minimum level of capacity at the start. This is a lesson that government has itself learned from the Dinaledi project, which was in many ways similar to the QLP[2] and which also had little success in a significant number of target schools: thus, the second phase of Dinaledi is being targeted at moderately functioning schools.

Conclusion

Standards based accountability consists of setting clearly defined expectations for pupil learning, and exerting pressure on schools, teachers and pupils to meet these standards. Around the world SBA has become the dominant mechanism adopted by politicians and administrators for attempting to improve the quality of schooling. However, as is the case with most major policy initiatives in education, its effects are mixed and often counter-intuitive, despite the best efforts of its proponents and detractors to paint it either as a simple solution to the problem of mediocre quality, or as the greatest enemy of progressive schooling. The South African evidence to date around SBA can be summarized in the form of five propositions.

Proposition 1: SBA, on its own, can lead to very significant changes in the indicators chosen to signal successful learning, and the higher the pressure on schools the greater the change.

What gets measured is what gets changed, and it is therefore important that the indicators are carefully chosen to reflect true measures of quality schooling. In South Africa there are two major problems with current standards. The first is that those

[2] The Dinaledi project, implemented in 102 schools nationally, achieved similar results to those shown in Table 4. However, one important difference between the two programmes is that, although Dinaledi was also targeted at very poor schools, it started from a significantly higher base than did the QLP. Thus, the frequency distribution of HG SC maths passes at the start of Dinaledi was skewed to the right when compared with the QLP distribution shown in Figure 3. Nevertheless, a very significant proportion of Dinaledi schools also showed no improvement as a result of the intervention (Taylor and Prinsloo, 2005).

defining the skills required by high school leavers are too low: this, together with poor school performance, is the root of the country's skill problem. Standards need to be moved progressively up the scale of cognitive demand, across the full range of school subjects, but most importantly for English Second Language.

The second problem with South Africa's standards is that they are defined, measured and valued at only one point in the system, the end of high school. Standards are best built incrementally, and clear signals need to be conveyed to and their achievement measured at all levels of the system. In order to address this problem, government's Systemic Evaluation process, which measures pupil performance in maths and language at the end of each school phase, needs to be improved and used for accountability and support purposes.

Proposition 2: There is a tendency to accommodate SBA pressure by the easiest means available, such as manipulation of the indicators, without necessarily improving learning.

The inclination to manipulation underlines the need to ensure that the intended standards are realized at all stages of the SBA process, including setting exam papers, marking pupil responses and moderating raw scores. Umalusi is in the process of limiting indicator manipulation in the first and third of these areas, but a major gap persists with respect to the marking of papers. This is a key node in the SBA process: it provides the opportunity to build capacity among teachers in the form of the fine-grained professional judgement required to recognize various manifestations of the standards; this is the first step towards replicating them in classrooms.

Proposition 3: There is inconclusive evidence to indicate whether SBA pressure, on its own, leads to improved school management, classroom teaching and pupil learning.

Strong evidence of manipulation in the period 1999-2003 has raised questions about the validity of indicators used to measure school quality in South Africa. As a result, the question of whether school performance improved over this period is difficult to answer. Circumstantial evidence indicates that some management slack was squeezed out of the high school system, resulting in institutions which are at least more efficient

in moving pupils along. However, it is unlikely that SBA has resulted in significantly better teaching, a task which will require considerable capacity building.

Proposition 4: A combination of SBA and capacity building aimed at strengthening internal accountability systems has significant effects on school performance.

Programme evaluations have begun to demonstrate that a number of school development projects in South Africa are achieving improved performance. Specific factors identified as key elements of the internal accountability systems required for effective teaching and learning include: time regulation, planning and monitoring curriculum coverage, and management of textbooks. However, this is work in progress and much ground needs to be covered before the complex processes of curriculum leadership and teaching are properly understood.

Proposition 5: In dysfunctional schools the combination of SBA and support has no effect.

The fundamental conditions conducive to effective learning do not exist in the majority of schools in South Africa. Such schools are impervious to any combination of SBA pressure and capacity building tried thus far. These schools need something else, and the international literature indicates that what is required is organizational development: fundamental issues such as removing ineffective school principals[3], mediating conflict, and building administrative capacity require attention before any learning is possible. Government has signaled its intention to move into this area, but this is an enormous task, given the weak state of the educational bureaucracy at all levels of the system.

References

J Allais E Davidson Y Reed (2004) **Report on English Senior Certificate Examinations**. Pretoria: Umalusi.

[3] In February 2006 both the national Minister and the Gauteng MEC for education threatened to close dysfunctional schools, but after protests from students, nothing has been heard of this issue since.

M Carnoy, R Elmore and L Siskin (2003a) **Introduction**, in M Carnoy, R Elmore and L Siskin (2003b)

M Carnoy, R Elmore and L Siskin (2003b) (eds) **The New Accountability: High Schools and High-Stakes Testing.** New York: RoutledgeFalmer.

I Carrim, C Julie and A Kitto **Report on an Investigation into Senior Certificate Mathematics Examinations of Various Years offered by the National Education Department and the Independent Examination Board.** Pretoria: Umalusi.

L Crouch (2005) South Africa and quality reforms: possible lessons. **Journal of Education for International Development, 1(1).**

L Crouch and T Fasih (2006) **Patterns in Educational Development: Implications for Further Efficiency Analysis.** Mimeo.

E Dempster, G Khumalo, M Mavovana (2004) **Umalusi Senior Certificate project: Biology group report.** Pretoria: Umalusi.

DoE (2004) **Education Statistics in South Africa at a Glance 2002**. Pretoria: Department of Education.

DoE (2005a) **Education Statistics in South Africa at a Glance 2003**. Pretoria: Department of Education.

DoE (2005b) **Education Statistics in South Africa at a Glance 2004**. Pretoria: Department of Education.

DoE (2005c) **Senior Certificate: Report on the 2005 examination results**. Pretoria: Department of Education.

R Elmore (2003) **Accountability and Capacity**, in M Carnoy, R Elmore and L Siskin (2003b)

Fast Facts (2001) Hold your applause for education. **Fast Facts, February 2001.** Johannesburg: South African Institute of Race Relations.

B Fleisch (forthcoming) Bureaucratic Accountability in the Education Action Zones of South Africa. To appear in **South African Journal of Education.**

Fleisch, B (2001) **Draft Research Report on the Education Action Zones, GDE.** Joint Education Trust: mimeo.

D Hopkins (2002) **The Aga Khan Foundation School Improvement Initiative: An International Change Perspective,** in Anderson, SE (ed) **Improving Schools Through Teacher Development: Case Studies of the Aga Khan Foundation Projects in East Africa.** Lisse, Netherlands: Swets and Zeitlinger.

D Hopkins, A Harris and D Jackson (1997). Understanding the school's capacity for development: growth states and strategies. **School Leadership and Management, 17(3)**,401-411.

HSRC (2005) **Improving Learning in South African Schools: Executive Summative Report to the Business Trust on the Quality Learning Project (QLP)**. Pretoria: Human Sciences Research Council.

P Kallaway, M Alexander and K Angier (2004) **Senior certificate research evaluation exercise: History report.** Pretoria: Umalusi.

D Koretz (2005) Alignment, high stakes and the inflation of test scores. **Uses and misuses of data for educational accountability.** National Society for the Study of Education: Yearbook 2005.

M Masemula, P Karsan and J. Clarke (2004) **An Investigation into the Senior Certificate Physical Science Examinations (1992-2003).** Pretoria: Umalusi.

MoE (2001). **Fourth Report to the President from the Minister of Education**. Pretoria: Ministry of Education.

S Nicols and D Berliner (2005) **The inevitable corruption of indicators and educators through high stakes testing.** Arizona State University: Education Policy Studies Laboratory.

J Muller (2000) **The sound and fury of international school reform: A critical review.** Joint Education Trust, mimeo.

N Taylor and C Prinsloo (2005) **The Quality Learning Project: Lessons for High School Improvement in South Africa.** Mimeo.

Umalusi (2004) **Have the standards of the Senior Certificate Examination declined? Summary report on the evaluation of the senior certificate examination.** Pretoria: Umalusi.

N Yeld, M Grobler and C Sekwane (2004) **Investigation into the examination of English Second/Additional Language of selected examination authorities in the years 1992, 1999, 2001 and 2003.** Pretoria: Umalusi.