Occasional Paper 1

August 2010

The South African national education evaluation system: What will make it work?

Godwin Khosa

JET EDUCATION SERVICES

August 2010

Table of contents

Background

The push for education evaluation ushered in by the Minister's National Education Evaluation and Development Unit Report of 2009 is a necessary development. This is so in spite of the doubts cast by critics, the minimal success of some earlier evaluation programmes (such as Whole School Evaluation (WSE) and the Integrated Quality Management System (IQMS) and the complexity of the task itself. A vehicle to evaluate the national education system is a key component of the broader strategy to improve education in this country. This paper therefore takes the opportunity to examine the overall national evaluation framework against the needs of the schools, the education system as whole and the aspirations of the society and poses questions about the potential that the evaluation framework offers for addressing the challenges we are facing. In order to answer the questions, the paper examines what these challenges are and how the national evaluation programme can be focused to overcome them.

It is, however, clear that while it is an important cog in the quality improvement machine, the inspection system the public has cried out for will not be a panacea or a quick fix for the problems of today's education system. Designing the evaluation vehicle is itself going to be a challenge. The purpose, objectives, coverage and methodologies of the evaluation system have to be carefully selected, as much as the processes of introducing and governing this function have to be responsibly undertaken. Some of these issues are described in the National Education Evaluation and Development Unit (NEEDU) report. This paper takes the NEEDU report as the starting point, draws lessons from other countries and JET's own research and development work and explores the education evaluation framework in South Africa with a view to outlining its key success factors and weaknesses. The exploration of the South African evaluation framework specifically focuses on the fundamental purposes, underlying theories and benefits of the various evaluation provisions, as well as the practicalities of its implementation.

Theoretical underpinnings of evaluation in education systems

The major education reform agenda over the past 30 years has been standards-based. In the main, it comprises student testing and public accountability (Carnoy et al, 2003: 2). While student testing programmes establish what learners know for diagnostic purposes, the accountability imperative, on the other hand, is purely concerned with establishing whether schools and education systems are creating the

expected value for money. There are several perspectives on and dimensions to the concept of accountability; these are explored in the following section.

Notions of national evaluation

The concepts of inspection, school reviews and evaluation are used interchangeably by researchers to refer to the process of determining the quality of education in schools. Generally speaking, national inspection and evaluation systems are concerned with improving education, learner achievement (or school outcomes) and institutional performance. However, their purposes, approaches and consequences differ from one country to the other. For instance, the Swedish system is concerned with auditing and assessment of both municipalities and schools, focusing on legal aspects (Skolverket, 2004:8.), while the United Kingdom is moving towards an increased focus on classroom observation of teaching and learning (OFSTED, 2004: 8.) The purposes of different evaluation systems include: establishing the value of the investment made in education (accountability); identifying areas that can be improved; benchmarking improvement; and establishing how the system works (research) and how it can be changed. Two approaches seem to dominate: a pure external evaluation approach that is carried out with minimal participation of the school management (e.g. in the United Kingdom); a complimentary evaluation approach based on internal audits of plans and targets set by schools and, in some cases, their municipalities (e.g. in Sweden and New Zealand).

Different national evaluation systems result in different consequences. Some result merely in criticisms directed to the relevant organisation and a requirement for action to be taken to rectify inadequacies; others result in serious sanctions such as closure of schools or sacking of principals and governing bodies (e.g. in New Zealand) (Thrupp, 1998).

Inspections are largely criticised for not taking into account socio-political contexts. Most of the critiques argue strongly that socio-economic factors are significant causes of schools' poor performance (Thrupp, 1998: 2). While this assertion holds some truth, the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SAQMEQ) study shows that some Southern African countries with better socio-economic contexts, such as South Africa, perform worse than countries with poorer socio-economic contexts, such as Mozambique (Moloi & Strauss, 2005). To support this view, reference can be made to an evaluation of the Khanyisa Education Improvement Project in Limpopo. Assessment data collected from 210 Limpopo primary schools, all of which are in a rural socio-economic region, revealed a marked difference in learner performance results among the schools in spite of their similar

socio-economic status: 95 of the schools achieved average learner performance gains in grade 3 literacy of ten percentage points (with the highest average gain score of 46 percentage points) over a two year period and 117 schools regressed by an average of 12.9 percentage points. This data shows that socio-economic status (SES) is not the key determinant of learner performance.

Additional criticisms leveled against inspections or external evaluations include the stress they cause for school staff, both at institutional and individual levels and the amount of time they take away from learning (an estimated two weeks out of 40 weeks of expected teaching) as schools prepare for and participate in the evaluation process. These specific criticisms signpost the need to make sure that evaluations have the desired effect and yield value in proportion to the time and effort they take.

Dimensions of evaluation and accountability

Describing the various evaluation approaches within the context of accountability, MacBeath (1999) posits that there are vertical and horizontal perspectives to evaluation. These perspectives recognise the fact that evaluation can be driven from different perspectives, depending on its motivation. Carnoy et al. (2003), on the other hand, have developed a three-dimensional theory of accountability, consisting of the internal-external, individual-collective and formal-informal dimensions. The theory is based on the premise that schools have conceptions of accountability embedded in the patterns of their day-to-day operations. According to this theory, accountability thus works best if the various forms of accountability (formal and informal) are aligned to individualised internal notions of accountability. In this way the accountability practices in organisations move from mechanical to organic accountability processes.

The focus of the systemic evaluation programmes

National evaluation systems have taken many different forms and adopted a variety of focus areas, depending on the perceived challenges facing the particular education system they serve. For this reason, it is important to carefully consider all the possible forms when a national system is designed and to ensure appropriate combinations and optimum cohesion of the various system components. There are two different approaches to determine an evaluation system's focus areas: the industrial production model and the hierarchical model. The industrial production model sees education as a production process involving school inputs, processes and outputs. On the other hand, the hierarchical model determines the focus of evaluation according to the levels of the system: learner level, teacher level, and school, regional and national level.

Since most evaluations are carried out for accountability purposes, they collect learner performance data as their primary data and use other data such as teacher qualifications and school resourcing to explain the patterns and profiles of learner performance. Therefore, the primary interest of such evaluations is in the output (learner performance), rather than the workings of the system. While such evaluations help to determine policy impact and value for educational investment, they unfortunately fail to help understand the organisational development or systemic processes. They are also unable to isolate the institutional human and social dynamics that contribute to the output and performance patterns which the evaluations diagnose. International educational assessments, an additional form of evaluation that national education systems are subject to, may also be characterised in this way. It can thus be said that most forms of educational evaluation focus on educational inputs and outputs and leave out the teaching and management processes.

While most national evaluations focus on institutions (schools), a unique form of evaluation is practiced in Chile: the evaluation of teachers' competence in the subjects they teach. The evaluation entails the production of a portfolio comprising: a written section; videotaped lessons; a supervisor's questionnaire; a peer interview; and a self-assessment questionnaire. 'The evaluated teachers receive qualitative reports detailing their reports on the different portfolio dimensions, describing their strengths and weaknesses including a final score' (Taut et al. 2010: 480). The reports are not made public but are used as the basis for rewarding and sanctioning teachers. High performing teachers are eligible for salary increases, while poor performing teachers have to undergo training and if repeatedly evaluated as 'unsatisfactory', lose their jobs. The evaluation is viewed by the education ministry and its subsidiaries as a way of advancing education quality and student achievement through the improvement of teacher competencies and performance. The evaluation outcomes are also used to identify teachers who could serve as teacher trainers and to better allocate professional development resources. Teacher unions are of the view that the assessment is good for professional development but should not be linked to student performance. As in the case of Chile, teacher evaluation is seen as a process which highlights the importance of professional teaching standards. In particular, Chilean teacher unions believe that, first and foremost, the evaluation process should trigger self-regulation and self-evaluation mechanisms among teachers.

How beneficial is education evaluation to schools?

The key concern of school evaluation initiatives should be how learners benefit from the process. There is overwhelming consensus that the primary goal of evaluation

is to raise standards. However, researchers and practitioners are divided on the extent to which schools benefit from evaluation, which forms of evaluation are most effective and how school evaluations should be conducted. As has been argued above, any form of evaluation works if it is aligned with the internal conditions of accountability. This means that there has to be internal institutional capacity to comprehend the need for accountability and to take cognisance of the performance gaps that are shown by the evaluation. Carnoy et al. (2003: 8) found that schools least aligned internally with accountability requirements are least likely to respond coherently to external accountability demands. Elmore (in Carnoy et al, 2003: 11) maintains that testing schools, rewarding and sanctioning them can take schools only so far in self-improvement; without internal coherence schools aren't able to make organisational changes. In this context, the benefit of an external evaluation, if it is strong, is to enforce standards. However, if the accountability system is weak, it both fails to enforce standards and to bring about school improvement from within. Partly in agreement with Carnoy et al. (2003), MacBeath (1999) argues for a bottom-up evaluation design wherein a school's self-evaluation precedes external evaluation. He contends that the bottom-up approach compensates for the weaknesses of top down approaches and their inability to pick up the real day to day experiences of learners and their teachers and the nuances that uniquely manifest in schools. There seems to be consensus that national evaluation systems for improvement should be built from the school level while national evaluation systems for accountability should be imposed externally from outside the schools – by districts, provinces and nationally. The question that arises from this observation though is, 'What is the best mix and sequence of evaluation approaches?'. Correctly put, MacBeath (1999: 1) purports that an unhealthy system relies on the constant routine of external evaluation. Following this line of thought, it therefore follows that in dysfunctional systems, the best way to approach the introduction of evaluation is to start with externally imposed evaluations with a view to enforcing standards and then gradually shifting the focus to more internal evaluations as a strategy to improve and sustain school quality.

In the end, the South African evaluation framework should be based on a thorough understanding of what is going on in schools compared with what the nation wants to see going on in schools. The evaluation system itself should be designed with the following principles in mind:

» The accountability and the learning roles of evaluation should complement each other (Stoop, 2008: 10). However, they should not necessarily be performed by the same institution.
» The goal of all evaluation should be to move accountability systems in schools

from mechanical to organic forms of accountability. This will entail building a locus of control and internally aligned sense of accountability within each school.

» All evaluation should help to improve institutional human and social dynamics that contribute to the differences in school performance patterns.

The South African education evaluation landscape

The education evaluation function in South Africa is regulated in terms of section 4 of the Education Act of 1996 which provides for the national minister to determine national policy for, inter alia, monitoring and evaluation of the well-being of the education system. This provision is given effect in a number of other education policies and programmes. Examples of these are the WSE, IQMS, Developmental Appraisal System (DAS) and Assessment Policy for the General Education and Training Band which provides for systemic evaluation to be conducted on a national representative sample of learners. These policies are analysed in detail in the NEEDU report. The 1997 amendment of the South African Schools Act strengthens the accountability provisions for education delivery. It requires the principals of schools to account for the academic achievements of the learners in their schools. Other important developments in respect of education evaluation come from outside the country, for instance, the international assessments that South Africa has participated in. Table 1 presents these international assessments and some of the key evaluation programmes that are active in the country.

Four of the seven programmes are international comparative assessments, while the other three are national programmes implemented at various levels of the system. All except the Developmental Appraisal System are sample based, meaning that they are not implemented in all the schools at the same time. The weakness with sample based evaluations is that they can hardly be used as accountability mechanisms for the system, since the results pertain to specific institutions. An alternative is to use rolling samples, aiming to cover all schools over a given cycle and maintaining focus on poor performing schools. The latter approach would be easy to implement in high schools where senior certificate results could be used to identify the poor performing schools.

As can be seen in Table 1, none of these programmes are comprehensive. Most collect and analyse data on school functionality and teacher profiles. Five collect and analyse learner performance and learner socio-economic status data. Almost none collect and analyse information on teacher knowledge and teaching practices. In fact, none of the programmes use observation as a method of data collection, meaning

that they do not have data on the social and human organisational dynamics that are responsible for the patterns of school outcomes.

Table 1: International and national evaluation initiatives in South Africa

	Institutional functionality	Teacher profiles	Teacher knowledge	Teaching practice	Learner performance	SES
TIMSS[1]	Yes	Yes	—	—	Yes	Yes
PIRLS[2]	Yes	Yes	—	—	Yes	Yes
SAQMEQ[3]	Yes	Yes	—	—	Yes	Yes
MLA[4]	Yes	Yes	_	_	Yes	Yes
Systemic evaluation	Yes	Yes	—	—	Yes	Yes
WSE/SSE[5]	Yes	—	—	Yes*	—	—
DAS[6]	—	Yes	—	—	—	—
	6	6	0	1	5	5

* Intended but never implemented
** IQMS was left out as it comprises WSE and DAS
1 Trends in International Mathematics and Science Study
2 Progress in International Reading Literacy Study
3 The Southern and Eastern Africa Consortium for Monitoring Educational Quality
4 Monitoring Learning Achievement Study
5 Whole School Evaluation/School Self-Evaluation
6 Developmental Appraisal System

The challenge in designing an evaluation system for South Africa is therefore to design an integrated evaluation system that covers key systemic elements and eliminates undue duplication in the collection of the required data, taking into account the key objectives of evaluating the South African education system.

The NEEDU report presents a significant opportunity to address the issues raised above. The report analysed the existing evaluation programmes and confirmed that most of the national evaluation programmes are flawed in that they are ambitious, complex and time consuming; assume that teachers have the requisite reflexive competencies; place little focus on the teaching and learning process and learner achievement; confuse the imperatives of the evaluation process (support vs. accountability); are not comprehensive in their focus areas; are weakly objective; and produce results that are sometimes questionable.

The NEEDU report itself can be criticised for failing to be specific and to prioritise

the functions of the evaluation unit. It proposes 11 functions for the unit, which is far too many. The categorisation of the NEEDU functions presented in Table 1 below shows that the recommended functions go against the very observation of the NEEDU report itself – that the existing programmes confuse the accountability and support imperatives of evaluation. The functions specified for NEEDU cover four imperatives: defining an accountability framework; reporting about the state of schools and factors responsible for school performance; developing capacity relating to internal evaluation; and generating knowledge to advise schools on school improvement and good practices.

Table 2: Functions of NEEDU

Evaluation Imperatives	Functions of the NEEDU
Accountability Framework	• minimum performance standards (2) • propose appropriate sanctions (6) • ensure coherence and complementarity of school evaluation (9)
Reporting	• state of schools (1) • account for the standards (3) • identify factors influencing school improvement (4) • make recommendations for redressing problems (5) • monitor school support (8)
Developmental	• strengthen the internal evaluation capacity in schools (7) • provide schools with evidence-based advice (10)
Knowledge generation	• evidence based advice on school improvement (10) • dissemination of good practice(11)

The numbers in brackets show the order in which the NEEDU report lists the functions.

The key functions of NEEDU should be to define the accountability framework and to report against it. The developmental and the knowledge generation imperatives are responsibilities of the provincial education districts and academic institutions respectively. The capacity of the district offices to support schools and the definition of the university research agenda need to be improved, rather than burdening NEEDU with developmental and research functions. Assigning these additional responsibilities to NEEDU risks dividing the focus of the unit and compromising its objectivity. With regard to the research function, it would be acceptable if it were specifically meant to provide internal support to NEEDU in its primary function of developing and maintaining the accountability framework, rather than servicing the national school quality research agenda.

The NEEDU report also does not make sufficient mention of the evaluation of teachers and teaching quality except for reference to the IQMS. In Chile, the teacher unions, government entities (local and regional) and universities concur that teacher evaluation improves teacher quality, competency and development, as well as the allocation of teaching resources. There is also consensus that teacher evaluation provides added benefits such as advocating professional teaching standards, providing feedback to teachers and encouraging dialogue among teachers about professional development (Taut et al: 2010). In South Africa, objective teacher evaluation has been the terrain of small scale research and project based initiatives. The benefit of these initiatives has largely been to demonstrate the potential of teacher evaluation, which teacher unions have vehemently opposed. The small scale initiatives also provide a space where evaluation protocols and methodologies can be tested jointly by researchers, education officials, teachers and their unions. These initiatives are also providing useful insights into teacher content knowledge. Research by Taylor and Moyana (2005) and Mabogoane and Pereira (2008) cautioned the nation about the low level of teacher content knowledge. More recent teacher evaluation conducted in the Northwest Province in the form of Rapid Based Assessment Testing (RBAT) in maths, science and language corroborates the earlier findings but also reveals some pockets of excellence in terms of teacher content knowledge. Table 3 below presents baseline scores from the RBAT administered to teachers from five high schools comprising a circuit in the Northwest Province. The tests were designed to cover a range of skills and content of the National Curriculum Statement (NCS) and were administered to teachers in the Further Education and Training (FET) phase who teach the subject tested.

Table 3: Teacher test scores on the RBAT in the Northwest Province

n= teachers	Mean	Highest Score	Lowest Score
Physical science (n=5)	54.0	77	35
Mathematics (n=8)	51.5	81	18
English (n = 8)	65.4	72	49

The RBAT results corroborate the findings of the earlier studies but specifically show that the majority of the teachers achieved scores in the region of 50–60. A positive aspect to note is that there are some teachers with an excellent command of the subjects they are teaching. On the other hand, as can be seen in the table, there are some teachers whose content knowledge is as low as 18% in maths and 35% in physical science. These teachers present the biggest challenge to school quality improvement. We may be tempted to conclude that the primary problem is that these teachers lack acceptable teacher content knowledge. Such conclusions are not correct in all cases.

Table 4 which presents the statistical profile of the schools, indicates that in this case the primary challenge is rather education resource planning. All the teachers tested in this case are from small schools that do not have a sufficient number of teachers to teach all the grades, let alone the different streams. In fact, the teacher who scored 18% in mathematics is a generalist teaching a number of subjects. This evidence concludes that in such small schools, there is no space for specialisation. Take the case of School 1, with five teachers having to teach all five high school grades and no less than seven streams.

Table 4: Learner to teacher ratio and teachers per grade

	Numbers		Teacher Ratios	
SCHOOL	**Learners**	**Teachers**	**Learners (L:T)**	**Grades (T:G)**
School 1	122	5	24.4	1
School 2	273	10	54.6	2
School 3	329	12	65.8	2.4
School 4	71	7	14.2	1.4
School 5	281	12	56.2	2.4
	215.2	**9.2**	**43.04**	**1.84**

L:T = Learner: teacher ratio; T:G = Teacher per grade

The teachers who failed the RBATs should either be retrained or, in most cases, be reallocated to the subjects they are best suited to teach; new and additional teachers should be provided to allow this to happen. What the above observations tells us is that the challenges faced by schools are multi-layered and complex. In order to understand them, they should be investigated in stages and the layers of school operations sequentially peeled off to gain an understanding of the organisational variables at play.

What is interesting about the design of the RBAT is that it signals to both the individual teachers and the officials responsible for teacher development the areas in which teachers have the least content knowledge, which teachers are excellent and which teachers don't have appropriate levels of content to be teaching particular subjects.

This form of evaluation thus presents useful information for professional development benchmarking, teacher allocation and teacher professional development needs of the school and the education system and an opportunity for individual teachers to set standards for themselves.

The challenge in implementing an in-depth methodology such as the RBAT approach, as opposed to traditional pen and pencil learner assessment, is that it requires high skills input and is a labour intensive and therefore costly exercise. University lecturers were used to design and implement the testing programme. An assessment report for each teacher was developed and shared with the relevant teacher and a composite report was drawn up for the benefit of the district and the project. While there is no question about the importance of this type of evaluation, the key questions that naturally follow are: how to rollout such a teacher evaluation programme in a system that comprises 400 000 teachers, how to target this large number of teachers and whether the country will have sufficient numbers of evaluators? It follows therefore that in-depth evaluations such as RBAT should be carried out in small samples following much larger profiling of school operations.

The state of school self-evaluation

It has been argued in the sections above that the existence of some internal school accountability system is a condition for a successful national evaluation system. Unfortunately, most South African schools seem to have weak internal capacity to conduct self-evaluations. In an attempt to improve school functionality, the Bojanala School Improvement Project in the Northwest Province initiated a process to help schools evaluate themselves by completing various questionnaires directed to school management, a sample of teachers and a sample of learners. Results from the first ten schools, four high schools and six primary schools, show that schools kept records for only 15.7 of the 47 areas covered by the school self-evaluation (SSE) instrument. High schools were found to have information for more of the monitored areas than primary schools. The number of areas for which schools collected information ranged from nine areas to 25. High schools had information for an average of 18.5 of the areas, compared to an average of 13.5 for primary schools. Table 5 shows information that was required of the schools.

Table 5: School information required and recorded

Information requirements	No. of schools that kept the information n = 10
Quarterly mathematics pass rate for Grade 12	0
Number of exercises in workbooks per term for Mathematics/Numeracy	0
Average number of learners writing at their correct level per term in Grade 3	1
Quarterly English pass rate for Grade 10	1
Average number of learners reading at their correct level per term in Grade 3	2
Average number of learners coping with mathematics for their level per term in Grade 3	2
Average number of teachers submitting planning per week	3
Ensuring that learners read, write and calculate daily	3
Average number of learners doing their homework without copying per week	3
Average number of teachers submitting assessments for moderation prior to giving assessments to learners per month	4
Stock register/inventory is always up to date	5
Average number of teachers absent per week	6
Average number of mathematics and language teachers who give learners at least one informal assessment per week	6
Average number of learners who are late at school per day	8
The financial statements for the previous year have been audited and presented to the SGB	8

The evidence from this project shows that most schools did not keep information relating to curriculum delivery. For 19 of the 20 curriculum related questions, information was available from only three or fewer schools. At least six out of ten schools could produce information about the 'Average number of mathematics and language teachers who give learners at least one informal assessment per week'. In a school effectiveness study involving 216 primary schools sampled nationally, it was established that 33% (71) of the schools had no information about teacher absenteeism. The same study found that, in the schools where registers were up to date, the frequency of absenteeism was high, indicating that there was no follow up on the information collected. In the schools that kept the relevant data (8.9%), the number of teachers absent on the day of the school visits was 159 compared to 33

in the 16.4% of the schools where the registers were not up to date (JET RNE Study, Unpublished). More data exist to suggest that school management teams know and monitor very little of what happens in schools. On the whole, these statistics show a general weak locus of control in schools. Specifically, they show:

» Weak curriculum leadership capacity, since the least information collected by management was on teaching and learning as opposed to generic information such as that on financial management and discipline;
» A lack of a culture of information use in schools. The schools did not keep the relevant education quality monitoring information; they did not identify improvement indicators, or set targets;
» Weak managers who were either ineffective or had no capacity to utilise the monitoring information, where it existed, for school improvement purposes. For instance, one school which had learner performance information categorised according to achievement targets was found to be oblivious of the declining learner performance between grades 3 and 6 and had no specified targets or plans to improve this dire situation.

It appears that, without key management information, school management is a rudderless endeavour. Furthermore, these observations raise the question whether such schools without a developed locus of control – an ability to organise internally, comprehend the complexity of schooling concerns and devise improvement strategies – would benefit from external evaluation. In other words, it is a question of whether the schools have the capacity to respond to an accountability oriented evaluation or grasp the performance expectations that it raises.

Conditions necessary for effective national education evaluation

The following recommendations are made towards making the national school evaluation effective:

1. A clear programme logic for the evaluation needs to be articulated

It is clear from the discussion above that the NEEDU report adopts the principles of a national evaluation approach but has not worked out the specifics. There is still a long way to go. First, the underlying theory of a South African evaluation programme

still needs to be clearly articulated. The programme's theory needs to be defined in order to delineate its envisaged consequences, principles, philosophy and the methodologies that will be shared by all involved. Bickman (1987), quoted in Taut et al (2010), defines a programme theory as 'the construction of a plausible and sensible model of how a program is supposed to work' (Taut et al, 2010: 478). Weiss (1973), quoted in the same source, purports that

> … the goal of social programs are often global, diffuse, diverse and inconsistent, vary over stakeholders and may have little to do with program functioning. One reason for this sorry state is that it often requires coalition support to secure adoption of a program. Holders of diverse values and different interests have to be won over, and therefore a host of realistic and unrealistic promises are made in the guise of program goals. (Taut et al, 2010: 479)

In addition, a programme theory may comprise two parts: one made up of ***hypothesised causal links*** and the other of a ***process map*** (Taut et al, 2010: 479). This should all be borne in mind when operationalising an ideal evaluation programme for South Africa.

2. The national evaluation should be designed to adopt a three-phase approach

Because schools experience different challenges, the national education evaluation system should be designed to take a phased approach that will allow universal profiling at first and then proceed to delve deeper into the dynamics of school operations. It should therefore comprise three phases: a ***universal testing*** of the schooling system to pick up cases of dysfunctionality; a ***light touch review*** of dysfunctional schools to determine how deeply seated the problems are – some of the changes may simply relate to resource provision and utilisation; and lastly, an ***in-depth review*** aimed at digging out deep seated educational challenges that may have to do with management operations, teacher and teaching quality and learner factors.

3. The rollout mechanisms of the evaluation system must separate the system's accountability and development imperatives

While improvement of the system requires both accountability and development inputs, the two imperatives should not be delivered through the same mechanism. NEEDU should be solely responsible for standards-based accountability and the

line management functions of the national and provincial departments (districts in particular) should take responsibility for development. Research and knowledge generation should also not be the primary responsibilities of NEEDU, but of academic institutions (universities and research councils). Too many roles given to NEEDU will weaken it and compromise its integrity.

4. The evaluation system should be strong and involve high-stakes testing

Research points out that for an external evaluation to have an effect in a weak schooling system, it has to be firm and based on the principles of high-stakes testing. The effect of such a system is to clarify to schools what the expected standards of achievement are. Therefore the evaluation methodology utilised and the teams carrying out the evaluations should be beyond reproach and the consequences of poor evaluation results should be serious.

5. The school support function should be strengthened

It is common knowledge that most districts don't have the requisite capacity to support schools. No model of district development has been adopted as yet and the district offices are largely underskilled. A national drive should be initiated to improve the capacity of the districts to address many of the development needs that a national evaluation would highlight, many of which are already commonly known. In addition to clarifying a district model, issues of district staffing and development programmes have to be urgently resolved to enable districts to function better.

6. Districts should be evaluated together with their schools

In most school evaluations, when fault is found, schools point to poor support from the district offices. Whether correct or not, these claims have to be analysed from the district level down. The best way to address this challenge is to include districts as an additional focus of the evaluation and to design the evaluation sample in such a way that a critical mass of schools in a district is sampled in order for the evaluation to pronounce on the effects of districts on school performance.

Conclusion

It has been demonstrated in the paper that the national evaluation system will not be a panacea for the complex education challenges in the country, nor will it be easy to design and rollout. It will most probably be mired in lack of cooperation, resourcing challenges and problems of extremely weak evaluation capacity in schools. These challenges, and many others not covered in this paper, necessitate the careful crafting of a clear programme theory for the national evaluation system, simple but tight methodological design, a fair level of differentiation in the implementation of the evaluation system in different schools and a series of sanctions that range from criticisms of the educational institutions to harsh sanctions for those that persistently fail to improve.

List of references

Carnoy, M., Elmore, R.F. & Siskin, L.S. (2003) *The new accountability: High schools and high stakes testing*. New York: RoutledgeFalmer.

Fuhrman, S. & Elmore, R.F. (2004) *Redesigning accountability systems for education*. New York: Teachers College Press.

JET Education Services (Unpublished) RNE study. Johannesburg: JET Education Services.

MacBeath, J.E.C. & National Union of Teachers (1999) *Schools must speak for themselves: The case for school self-evaluation*. London: Routledge.

Mabogoane, T. & Pereira, C. (2008) Using trends to hold schools accountable. Presentation to the Association for Evaluation and Assessment in South Africa Conference, Pretoria, 2008.

Ministerial Committee on a National Education Evaluation and Development Unit (NEEDU) (2009) Final report. *Government Gazette* 2133, 17 April 2009. Pretoria: Government Printer.

Moloi, M. & Strauss, J. (2005) *The SACMEQ II Project in South Africa: A study of the conditions of schooling and the quality of education*. Harare: SACMEQ.

Office for Standards in Education (OFSTED) (2004) The future of inspection: A consultation paper. London: OFSTED. [Available from http://www.ofsted.gov.uk/Ofsted-home/Forms-and-guidance/Browse-all-by/Other/General/The-future-of-inspection-A-consultation-paper]

Skolverket/Swedish National Agency for Education (2004) Inspecting for improvement – A brochure about the National Agency for Education's Educational Inspectorate. Stockholm: Skolverket. [Available from http://www.skolverket.se/publikationer?id=1358]

Stoop, G. (2008) Increasing evaluation capacity in the education sector. *New Zealand Principals' Federation Magazine*, June 2008: 9–10.

Taut, S., Santelices, V., Araya C. & Manzi, J. (2010) Theory underlying a national teacher evaluation program. *Evaluation and Program Planning*, 33(4): 477–86.

Taylor, N. & Moyana, J. (2005) Khanyisa Education Support Programme: Baseline study part 1: communities, schools and classrooms. Johannesburg: JET Education Services.

Thrupp, M. (1998) Exploring the politics of blame: School inspection and its contestation in New Zealand and England. *Comparative Education*, 34(2): 195–209.

Published in 2010 by
JET Education Services
5th Floor Forum 1
Braampark
33 Hoofd Street
Braamfontein
2001

www.jet.org.za

Design and lay-out by COMPRESS.dsl