

Assessment systems and practices: Implications for education policy

**Keynote address:
Global Conference on Education Research
Prague, April 2005**

**Anil Kanjee
Assessment Technology and Education Evaluation
South Africa**

akanjee@hsrc.ac.za



Introduction: Quick Test



Purpose

- **FOCUS**: Use of assessment, and its impact on policy processes, for improving the quality of education.
- **EMPHASIS**: Systems and practical level
- **APPROACH**: identify key lessons for using assessment to effect education reform.
 - Development of Assessment Systems
 - Practical implications at classroom and national level



Outline

- **Current context of assessment**
- **Definition and use**
- **Key lessons: Classroom, National, International**
- **AQEE to improve learning model**
- **Systemic Validity**
- **Strategies for moving forward**



Criticality of assessment

- **setting of education targets**
- **emphasis on a 'results agenda'**
- **greater focus on accountability**
- **value of evidence based decisions**
- **effect of globalisation**
- **recognition of role of assessment in improving learning and teaching**



Education Policy Processes

- Central to education development and improvement of the system
- Policy processes – policy making, implementation, analysis & evaluation
- Implicit in any policy – Cause and Effect
- Require valid and timely information to effect relevant policy processes
- Policy analytic triangle Is the policy: (Bhola, 2000)
 - principled,
 - professionally sound
 - practical



Definition of terms

- **Measurement:** the process by which a value, usually numerical, is assigned to the attributes or dimensions of some concept or physical object
- **Testing:** the process of measuring one or more concepts, usually under standardized conditions
- **Evaluation:** the process of arriving at a judgment about abstract entities such as systems, structures, programs, curricula and organizations
- **Assessment:** “The process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his or her progress, strengths and weaknesses, to judge instructional effectiveness and curricular adequacy and to inform policy (AFT, NCME, NEA, 1990, p.1)”, and usually involves a range of different qualitative and quantitative techniques



Uses of Assessment

- **Learner Perspective**
 - **Choose**
 - **Learn**
 - **Qualify**

- **Decision Maker Perspective**
 - **Select**
 - **Monitor & Evaluate**
 - **Hold Accountable**



Focus of Assessment

- **Cognitive**
- **Skills**
- **Behavior**
- **Attitude**
- **Value**



Types of Assessments

- **Assessment of Learners**
 - Classroom
 - National (District)
 - International / Regional
- **Assessment of School staff**
 - Teachers
 - Principals
- **Assessment of Education officials**
 - District
 - Provincial/Regional
 - National



Development of Assessment Systems



Key lessons: Brazil

- **Assessment system developed to “co-ordinate the development of educational assessment systems and organize the information and statistic systems to assist policymakers at all level of the system in making appropriate decisions”** (Guimaraes de Castro, 2001b)
- **Key features**
 - single federal agency, National Institute for Educational Studies and Research (INEP) to develop and coordinate the country’s national assessment system
 - linking of the country’s assessment and information systems for use in monitoring policy formulation and implementation
 - inclusion of the higher education sector (undergraduate) as a component of the national assessment system,
 - Multi-level regular assessments – secondary, national, international



Key lessons: Chile

- Regards the assessment system in Chile as one of the most comprehensive and best-managed assessment systems in Latin America, and one that has served as a strong tool for implementing required reforms and led to increasing learning Wolff (1998)
- **incomplete**



Key lessons: Chile

■ Key Features

- use of census sampling
- The gradual improvement of the assessments over time. For example, delivery of school reports.
- The increased use of assessment results - intervention programs for low performing schools, allocation of financial rewards to schools
- Massive media campaigns directed at teachers, principals and parents regarding the purpose and use of national assessments.
- A comprehensive strategy for disseminating results and suggestions for teachers on how to improve learner performance - relevant manuals, video tapes, trained supervisors
- The assessment of the affective domain, i.e. self-concept, attitudes towards learning, peer and social relations, vocational orientation and values acquisition..
- The categorization of schools by the typical socioeconomic level and reporting of schools results within the levels.
- Participation in regional and international assessments

■ Integrated Curriculum, Assessment & Research division in MoE



Key lessons: Jordan

- **Revision of the General Secondary Education Certificate Examination (Grade 12) for the improvement of assessment procedures, as well as the skills and expertise of teachers.**
- **Features:**
 - The use of external expertise for technical assistance and training in developing the assessment system,
 - The focus on classroom assessment and the learning and teaching process in the form of diagnostic testing and implementation of the “investigation” approach
 - The emphasis on the vocational education sector as an integral part of the system
 - Participation in regional and international studies including: TIMSS (1999 & 2003), IAEP II (1991)



Key lessons: Mauritius

- **Changes in assessment - result of changes to education system**
- **Highlights:**
 - The use of examination results to identify weaknesses in learner performance and provide information to promote relevant reform
 - The use of both national and international examinations to certify learners' completion of formal schooling.
 - The identification of Essential and Desired Learning Competencies for improving the learning and teaching process.
 - The provision of “standardized” diagnostic instruments for numeracy and literacy to assist teachers in detecting learning difficulties.
 - Participation in regional studies - SACMEQ (1998, 2002) & MLA (1994, 1999) & use of information to inform policy



Key Lessons: Netherlands

- **Inspectorate of Education supervises education and assess school quality**
- **Features:**
 - Evaluate quality of individual schools from context, input, process and output perspective
 - System comprised of school self evaluation, national examinations, national and international assessments
 - Pupil Monitoring System – annually for 8 years, census
 - Track performance of individual learners over time
 - Software to assist teachers in analysis and recording
 - Conducted by external agency



Key lessons: New Zealand

- **National assessment (NEMP) – relevant info for policy makers + professional development of teachers**
 - Used wide variety of data in NEMP, observations, group work, practical work,
 - Massive support – ARBs,
 - Asttle – using NA + classroom teacher support,
- **incomplete**



Key lessons: South Africa

- **Assessment system implemented as part of complete “overhaul” of apartheid education**
- **Highlights**
 - National Qualifications Framework - education and training
 - Standards and outcomes for learners + Adults
 - Certification end of Compulsory education Grade 9 – CTAs
 - Revised Matriculation Exam - FETC
 - Whole School Evaluation + Systemic Evaluation
 - “Poverty index” for each school – decision making
- **Current reform relatively new – difficult to judge success**
 - **Minimum attention on assessed – Revised NEW curriculum after 2 years**



Key Lessons: Uganda

- **Introduced a Statistics, Monitoring and Evaluation Division within Education Planning Department of Ministry**
- **Features:**
 - Comprehensive framework that links monitoring, evaluation, reporting, data collection & research for MoE
 - Integrated into EMIS of MoE
 - Specifies key indicators for entire education sector
 - Address funding issue
 - Provides reporting framework



Implementing Assessment practices



Lessons & Challenges: Classroom

- Using LSA to judge schools



Lessons & Challenges: National



Why emphasis on national assessments?

- Potential to focus on entire education system
- Significant increase in last decade, especially in Developing Nations
- Use of NA likely to increase in future
- Summative and formative role – policy processes and classroom practices
- Cost effective way of obtaining information
- Generally low stakes
- Greater utility value compared to international assessments



Relevance for developing nations?

- Effective use of NA – significant potential to improve education systems – many of which are in state of transition
- Range of national “monitoring” practices – which might not be reliable & valid
- Limited discourse and research to address specific needs of developing nations
- Specific interest in the use of assessment in developing countries –
- Look at different practices/approaches between developed and developing nations
- Identify criteria for decisions regarding NA and International Assessments



Overview - NA in Developing Nations

Examinations or other proxy indicators	Regional testing or international agency	National monitoring program
Botswana; Cape Verde; Comoros Islands; Eritrea; Gambia; Guinea; Mauritius; Togo; Tanzania (Zanzibar); Zimbabwe	Kenya; Mali; Malawi; Mauritius; Mozambique; Namibia; Nigeria; Seychelles; South Africa; Swaziland; Tanzania (Mainland); Tanzania (Zanzibar); Uganda; Zambia; Zanzibar; Zimbabwe	Benin*; Congo*; Gambia; Lesotho; Madagascar*; Zaire; (Congo Democratic Republic)*; Zambia
Algeria; Djibouti; Gaza Strip; Iran; Iraq; Jordan; Morocco; Oman; Saudi Arabia; Sudan; Tunisia; Yemen	Oman	Jordan; Lebanon*; Morocco*; Syria*; Tunisia
Bhutan; Maldives; Nepal; Sri Lanka	Bangladesh; Maldives; Nepal; Pakistan; Sri Lanka	
Cambodia; China; Fiji; Indonesia; Laos; People's Democratic Republic; Papua New Guinea; Malaysia; Vietnam	Cambodia; China; East Timor; Mongolia	Philippines; Thailand
Barbados; Belize; Ecuador; Guyana; Haiti; Nicaragua; Panama	Argentina; Bolivia; Brazil; Chile; Colombia; Costa Rica; Cuba; Dominican Republic; Honduras; Mexico; Paraguay; Peru; Venezuela	Brazil; Chile; Colombia; Costa Rica; El Salvador; Guatemala; Honduras; Jamaica; Mexico; Nicaragua; Paraguay; Peru; Uruguay; Venezuela
Armenia; Azerbaijan; Bosnia & Herzegovina; Croatia; Georgia; Kazakhstan	Croatia	Mongolia; Romania (under review)

NA – Issues to address

- Value of, and need for, NA accepted
- However different understanding and purposes
- Key question: How can NA be used DIRECTLY support for the teaching and learning process?
- What designs and methods are required that both sustainable within the system and provide relevant, reliable and valid information?
- One problem – different understanding of the use and purpose of NA



National Assessment: Purposes

- Evaluate and monitor policy
- Determine levels of learner achievement
- Identify areas in need of intervention
- Allocation of resources
- Monitor standards
- Evaluate the quality of education
- Increase public awareness
- Inform political debate
- Determine learner strengths and weaknesses
- Monitor and evaluate intervention programs
- Evaluate the effectiveness of the curriculum
- Compare student performance
- Identify students with special needs



Definitions of NA

- Range of interpretations and different conceptions
- General sense of “confusion”



Definitions – examples 1

- “systematically measure typical levels of **learner achievement** in relation to **national standards**” (Crighton, J, 2002, p. 4). UNICEF Education Update
- “systemic regular measures of learning achievement in a country designed to assist **policy making**” (Lockheed & Murphy, 1996, p. 2)
- “regular and systematic measurement exercise designed to determine what students have learned as a result of their educational experiences” (UNESCO, 2000, p 14)
- “[a] nationally representative and continuing assessment of what America's students know and can do in various subject areas” (NAGB, 2004)



Definitions – examples 2

- “a process in which various measures are used to gather data for the purpose of **monitoring school systems**” (Capper, 1996, p. 30)
- “an exercise designed to describe the **levels of achievements**, not of individual students, but **of a whole education system**, or a clearly defined part of it” (Kellaghan & Greaney, 2001, p. 33)
- “**assessment materials** designed to be used by teachers in Scottish schools to confirm their judgments about pupils’ levels of attainment in English language (reading and writing) and mathematics 5-14”. (Scottish Executive http://aifl-na.net/na/general_intro.aspx Retrieved 17 February 2004)



Possible explanation

- Used for range of reasons with very similar/exact designs and methodologies for collecting, analyzing, reporting and disseminating information
- Use and interpretations of evaluation vs assessment vs monitoring
- Tendency to regard all forms of LSA as similar when in practice these are VERY different and serve different purposes



Require new name?

- I do think we require a new name and definition that would eliminate ambiguity
- lead to a “more” universal understanding – especially between researchers/academics on the one hand and policy makers, teachers, principals, DoE officials on the other hand
- Focus on the essential aspect of education – improving learning (and teaching)
- New Zealand – good model

National Education Monitoring Project



More suitable definition?

- the process of obtaining relevant information from the **entire education system** to **monitor and evaluate** the performance of **learners and other significant roleplayers** as well as the **functioning of relevant structures and programs** within the system for the purpose of **improving learning**
- Role of policy makers and teachers the same – improve learning – although functioning at different levels



Definition – essential criteria

- for *improving learning*
- from the *entire* education system
- to *monitor and evaluate*
- performance of *learners* (ESSENTIAL CRITERIA)
and other roleplayers
- functioning of *structures and programs*

- deliberate in its omission of the process by which these studies should be conducted and the specifics regarding how information should be used



Data collection and Analysis: Similarities!

	National Assessment	Accountability Assessment
Stakes	Low for participants High for MoE	High for participants Low/High for MoE
Target population	Specific grade (s)	
Selection of participants	Census or sampling	Census
Contents assessed	Single or multiple content areas	
Content coverage	Limited or comprehensive	Limited
Item formats	Subjective, objective or both	
Number of test forms	Single or Multiple	Single
Administration process	External agency, MoE or both	
Frequency	Varies according to objectives	



Reporting, dissemination & utilization of results

- Most instances - reports are meant for policy makers (used by researchers)
- Generally – context of education, learner performance, (sometimes) factors affecting LP
- Rarely – information pertaining to teachers
- Some exceptions – e.g. Chile (census approach)
- Limited information on the use of NA and its impact on education practice



NA in developing countries

- Many countries use proxy measures – e.g. examination results (by design or unintentionally)
- Many countries “introduced” NA via participation in regional (or international) studies – mainly conducted/promoted by UNESCO: MLA, SACMEQ, PASEQ
- A small number of countries have or have begun establishing their own regular NA
- REASONS: vary from high cost to limited expertise to lack of political will, etc



Using NA for L&T: Teacher

FOCUS: On current studies & available data

Many proposed uses for L&T process.

- Mechanisms for PD
- Identifying student strengths & weaknesses
(curricular sub-domain; cognitive process, sub-groups)
- Proposing possible solutions to specific problems
- Provide items for use in classroom
- Exemplars for good assessment practice
- Highlight critical skills that students should master
- Report on whether standards are being met
- Evaluate policy implementation

However – to date limited information on how useful/successful these have been



Pre-requisites

Need in-depth understanding of:

- specific information needs of teachers that can be addressed by NA
- reporting practices, and associated analysis & collection of data, that would support teachers to improve classroom practice
- how teachers understand and translate information into practice (Brown, 2003; Chapman & Snyder, 2002)
- how students interact with assessment information (Black & William, etc.)



Example: issues in SA

- Teacher needs differ by level of experience and school type (location)

e.g. – identify problems areas vs likely solutions

“we don’t want information on what problems students are experiencing. We know what problems our students have. What we don’t know is how best to address these problems”
(Grade 3 teacher – urban school, November 2003)

- Different uses of information from NA

e.g. – use of items in class



Methodologies & Designs

- Latest & most sophisticated methodologies vs Cost effective and “good enough”
- Depends on national context and purpose
- **QUESTION: Do we really need sophisticated methods and designs to evaluate the following schools?**



Note however
We do have many very good schools
with lots of learners have







DISHI
P.H.







21 8 2002

I went to her
house.
A bicycle has
two wheels.
She has two
red pens.
Tom came to
my school.



1. She was at home.
2. They were in the
bus.
3. Dad was at the
shop.
4. We were coming
to school.

1. He has a book.
2. We have a new
car.
3. The children have
many toys.
4. Mum has a red
dress.





21 8 2002

Lessons & Challenges: Classroom



General Trend

- Greater recognition of value of classroom assessment in effecting change
- Provision of information from National Assessments to teachers
- Preference for National Assessments that meet both needs of policy makers as well as teachers
-



Lessons & Challenges: International



Issues in IA

- Can IA support learning and how?
- Issues of curriculum relevance, comparability,



Lessons & Challenges: School staff

- **Development Appraisal System in S.A.**
- **Incomplete**
- **Perhaps leave this out and note as Limitations**
- **i.e. Assessment of School staff and district officials.**
- **Per**



Lessons & Challenges: Education officials

- **Perhaps leave this out and note as Limitations**



Lessons & Challenges: General

- **Methodology: Best vs Sufficient**
- **Greater and more effective use of available information**
- **Use of information to DIRECTLY support improvement of learning and teaching process**
- **Incorporating classroom based assessments into LSA programs**
- **Use of ICT vs Resources**



AQEE to improve learning

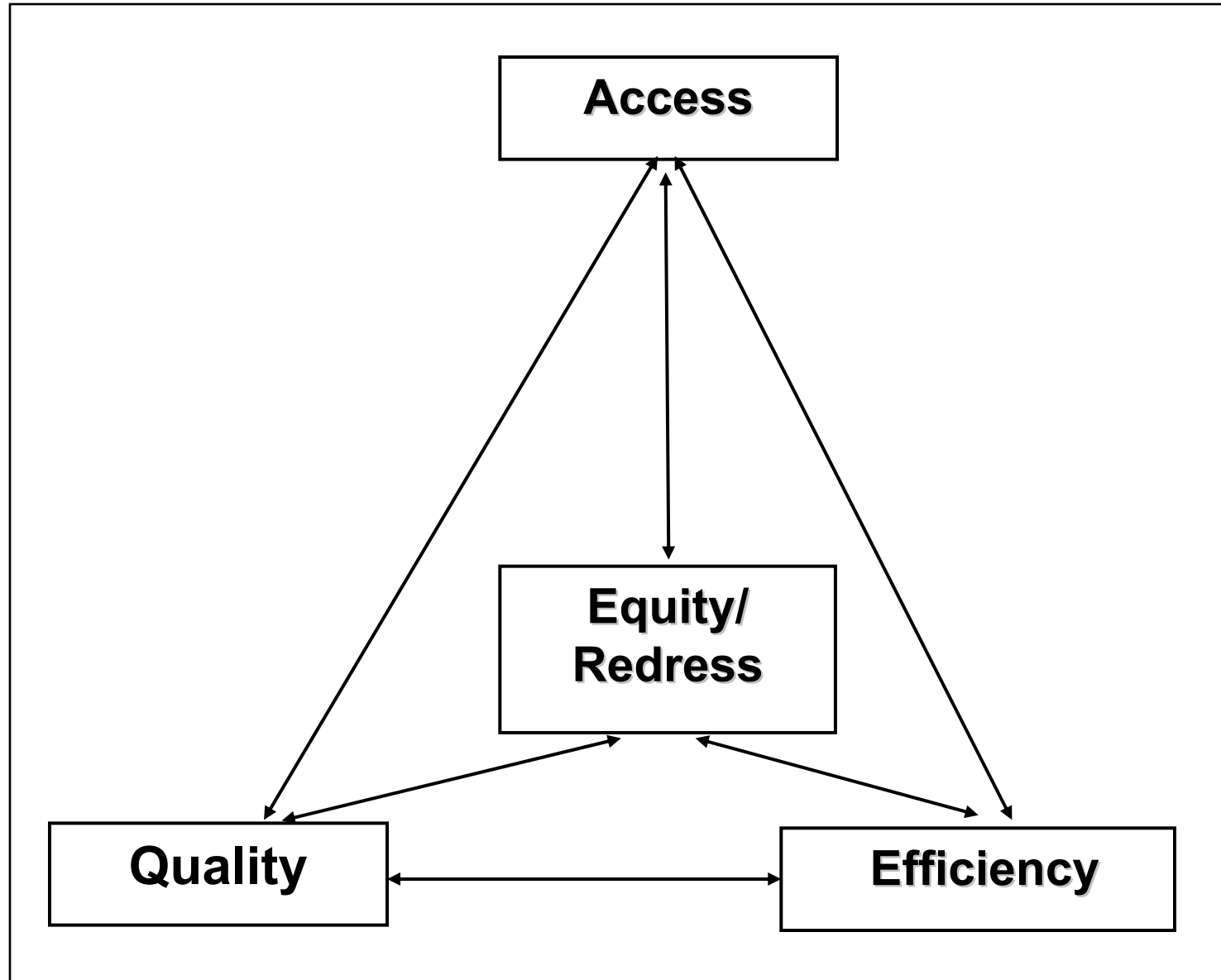


AQEE Model

- **Model based on systems approach**
- **Allows for defining & understanding key policy goals that define Access, Quality, Efficiency & Equity/Redress**
- **Applicable at different levels of education system: national, provincial, district, school**
- **Each policy goal viewed as specific framework defined by indicators and indices**
- **Thus adaptable to any education system at any level**
- **Framework for evaluating education systems**
- **Currently being piloted in SA**



AQEE to improve learning



AQEE “Policy Goals”

■ Access

- Getting to school
- Getting into school
- Getting through school

■ Quality

- What learners should know
- Where learning occurs
- How learning takes place
- What is actually learnt

■ Efficiency

- Functioning of structures and systems
- Availability, allocation and use of human & physical resources
- Throughput and repetition rates

■ Equity/Redress

- Inclusivity
- Absence of unfair discrimination



Systemic Validity

A systemically valid test is “... one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure.”

Frederiksen & Collins (1989)

Assessment practices and systems are systemically valid if they generate useful information that supports the (continuous) improvement in one or more aspects of AQEE within the education system, without causing undue deterioration in other aspects or at other levels.

Braun & Kanjee (2004)



Example: Selection Exams

- **Well constructed exams that serve the intended purpose of selecting students for the next stage**
(Heyneman and Ransom, 1990)
- **Meet the Fredriksen & Collins criteria**
- **Examination practices have deleterious effects on the education system**
(Kellaghan & Greaney, 1992; Govinda, 1988)
- **Fail the Braun & Kanjee criteria**



Assessment & Education Reform – Impact on Policy



Context: Education Landscape

- **Substantial disparities in opportunity to learn and in achievement**
- **Low levels of achievement relative to global standards**
- **Limited human and physical resources**
- **Inefficiencies in allocation and use of resources**
- **Other socio-cultural and political challenges**



Assessment paradox

- **Low stakes assessment provides useful information for improving education practice - no incentives for reform**
- **High stakes assessment - linked to specific incentives for reform - many unintended consequences**



Conflicting views

- **Assessment is a fair and objective way to set and maintain standards, to spearhead reform at both the levels of policy and practice, and to establish a basis for meaningful accountability.**
- **Assessment an instrument for maintaining the status quo, grossly unfair and educationally unproductive**



Systemic validity & AQEE

- **AQEE framework provides the basis for judging how education systems function in support of societal needs**
- **Use systemic validity for prospective evaluations of proposed reforms against AQEE**
- **Should be complemented by cost-benefit analysis**



Attaining Systemic Validity?

- **Systemic validity an ideal to be approached**
- **Prerequisites**
 - **Clarity**
 - **Coherence**
 - **Consistency**



Successful approaches **INCLUDE??**

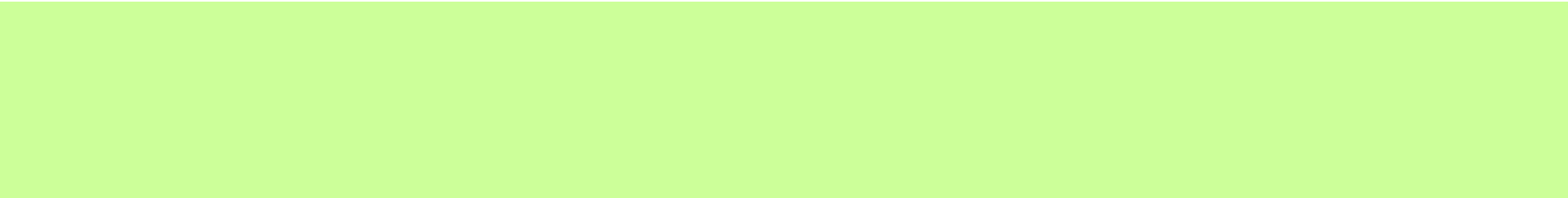
Brazil

- Single federal agency to coordinate entire national assessment system (including higher education)
- Linking of EMIS system for Monitoring & Evaluation

Jordan

- The use of external expertise in developing the assessment system (K – 12)
- Specific focus on learning and teaching based on introduction of diagnostic assessments and the “investigation” approach
- **INCOMPELTE**





Concerns: Assessment for education reform

- **Failure to identify relevant skills and competencies to function effectively in society**
- **Technical and substantive aspects – e.g. quality of data, language differences, inequities in opportunity to learn**
- **Reporting and dissemination**
- **Unintended consequences**



Constraints

- **Current state of education system and assessment sub-system**
- **Limited expertise and infrastructure**
- **Political will/champions**
- **Cost**



Strategies for moving forward - 1

- **Models of school improvement that make explicit the links between assessment, curriculum & instruction**
- **Cultivate among all stakeholders an appreciation of the power and cost-effectiveness of assessment (at different levels)**



Strategies for moving forward – 2

- **Model of delivery - use of assessment related activities in teacher professional development as a mechanism for better content knowledge and for generating more useful feedback**
- **Example – use of Assessment Resource Banks in South Africa**



Assessment need to include ABET

- Point about greater focus on ABET given overwhelming evidence that children of educated parents get higher scores.
- Also MLA results, learners who had someone at home to assist with school work got higher scores



Assessment

- To assess attitudes, values, skills & behaviors as well



Strategies for moving forward – 3

Leveraging global resources

- 1. Full participation in regional and international studies**
- 2. Participation as “associate members”**
- 3. Harness existing networks, e.g. AMASA**
- 4. Countries can use one or more of these mechanisms to draw available information on: frameworks, items, scoring rubrics, sampling designs, pedagogy, etc.**



LIMITATIONS

- **Have not adequately addressed Lesson & Challenges pertaining to the assessment of schools staff and education department officials**
- **Have not looked at assessments that go beyond cognitive level – i.e. attitude, values,**



Conclusion - 1

Assessment is an under-utilized and cost-effective tool in the drive for educational improvement. It can serve to enhance all aspects of AQEE, either directly or indirectly.



Conclusion – 2

Acknowledge that

“Testing alone cannot improve learning, nor can it necessarily make education systems more responsive. But it does tune societies and governments alike to the possibilities of their schools and education systems. And, if the past is any guide to the future, well-designed and applied assessments can change the course of education reform and the menu inputs used to promote it” (Schiefelbein & Schiefelbein, 2003, p. 154).



Final Note:

- In my preparation, I am certain that there are experiences and information that I left out or did not accurately present.
- I would thus appreciate any comments, suggestions, additions or information that would improve on this paper

akanjee@hsrc.ac.za

akanjee@telkomsa.net

