

Whither National Assessments?

The Nexus of Policy Intentions and Implementation

16-17 April 2013

Qetelo Moloji

Presentation Outline

- Introduction
- Definition of national assessments
- Overview of the landscape of national assessments
- Policy imperatives for national assessments
 - Purpose(s)
 - Reporting of results
 - Utilisation of assessment data
 - Impact
- Theory of Action and Policy: Recommendations

Introduction

- Interest in large-scale national assessments of learner performance is increasing across developed and developing countries
- South Africa has weighed in, not only into international assessments but has also conceptualised and implemented a unique approach to national assessments.
- South Africa's approach to national assessments is unique both in scale and intended purpose(s)
- There's need to ensure effectiveness, sustainability and quality of the assessments based on well-informed policy intentions and policy implementation

Definition of national assessments

- **National assessments** → large-scale assessments (national and international) of learners in identified subjects and grades, conducted annually or periodically on populations or samples of learners to generate data that must inform decision-making
- Distinction between national assessments and examinations → often initially stated as distinguished by “low- and high-stakes” implications, respectively, but often the distinctions end up being **obfuscated**

Focal questions steering literature review

- The often-observed **obfuscation** at the *nexus* of policy intentions and policy implementation regarding national assessments provided the impetus for writing this paper
 - Are there common reasons shared by countries that conduct large-scale assessments?
 - What distinguishes education systems that have realised sustained and successful large-scale assessments from those that did not?
- What lessons, if any, can we learn?

Views against national assessment ...

- supports policy-decisions that have not been adequately scrutinized and are based on measures with important psychometric limitations
- acts as a form of educational and social control that subverts the professional autonomy of educators
- holds teachers responsible for results with inequitable resources and contributes to feelings of lower self-efficacy
- undermines the quality of education by penalizing divergent thinking, creativity, and intellectual work in general

Views against national assessment ...

- widens the gap between minorities and majority students by ignoring key factors known to affect student performance such as socio-economic status, language of origin, and the students' physical and/or emotional health
- narrows and distorts the curriculum by encouraging "teaching to the test" techniques which take valuable time away from non-tested subjects, particularly when high-stakes are attached to results
- provides information that is of relatively little use to improving classroom practice

Views in support of national assessments ...

- Make students work harder and perform better on national and international assessments
- identify the most successful teaching practices and proficient teachers
- enhance teachers' reflective and critical thinking when planning instruction
- lead to positive increases in educators' knowledge about testing and of testing issues
- improve teachers' assessment and instructional practices, particularly when they are involved in marking these assessments

Views in support of national assessments ...

- stimulate action research that serves as a basis for school improvement
- provide schools with valuable information about the consequences of their past practices and program effectiveness
- damages children's self-concept and leads to student disengagement

Overview: Key observations

- National assessments have often abounded following abandonment of the now unpopular school inspectorate systems, apparently to fill the monitoring “void” left by the transformation
- Uptake of national assessments has seen unprecedented increases following the two EFA world conferences (2000 and 2005) in Thailand and Dakar, respectively
- Key message in Dakar was: “Ensure that all learners achieve measurable learning outcomes of a high quality”, has increase in national and international and regional assessments

Key observations: Purpose of assessment

- Ranking, medium-stakes, often followed by either stick or carrot
- Accountability , high-stakes, name-and-shame
- Value added, low- to medium-stakes, recognises impact of contextual factors and rewards effort
- Curriculum review, no-stakes, recognition of importance of evidence in decision-making
- Monitoring (access, quality, equity and efficiency), no- to low-stakes, preferred where there are socio-economic disparities
- Set and maintain standards, low- to medium-stakes

Can We Have a One-Size-Fits-All Assessment?



Some of unintended consequences

- Education systems tend to fall easily into the temptation of either changing original purposes of assessments or adding (overloading) new demands
- Unplanned change of purpose of assessments resulting in resistance
- Under- or non-utilisation of assessment data resulting in low returns on incurred costs
- Teachers teaching to the test

Approach to reporting results

- Reporting style influences uptake and utilisation of assessment results
- Reporting results in raw scores (percentages) is fast-losing popularity because of evident shortcomings → no-information, often norm-referenced
- Latest “theory-laden-measurement” approaches to assessment (IRT/Rasch) enhance opportunity for meaningful reporting and increase prospects for use of assessment data → information-rich, criterion-referenced
- Use of performance standards models expected levels of performance → everyone gets to know “how good is good enough”

Setting and reporting standards

Logit

3.00

2.75

2.50

2.25

2.0

1.75

1.5

1.25

1.0

0.75

0.5

0.25

0.00

-0.25

-0.5

-0.75

-1.00

-1.25

-1.50

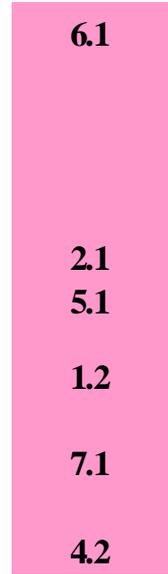
-1.75

-2.00

-2.25



LEVEL 1



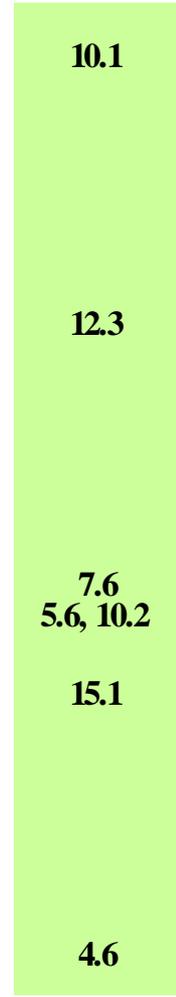
LEVEL 2



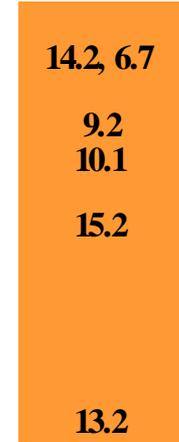
LEVEL 3



LEVEL 4



EXEMPTED





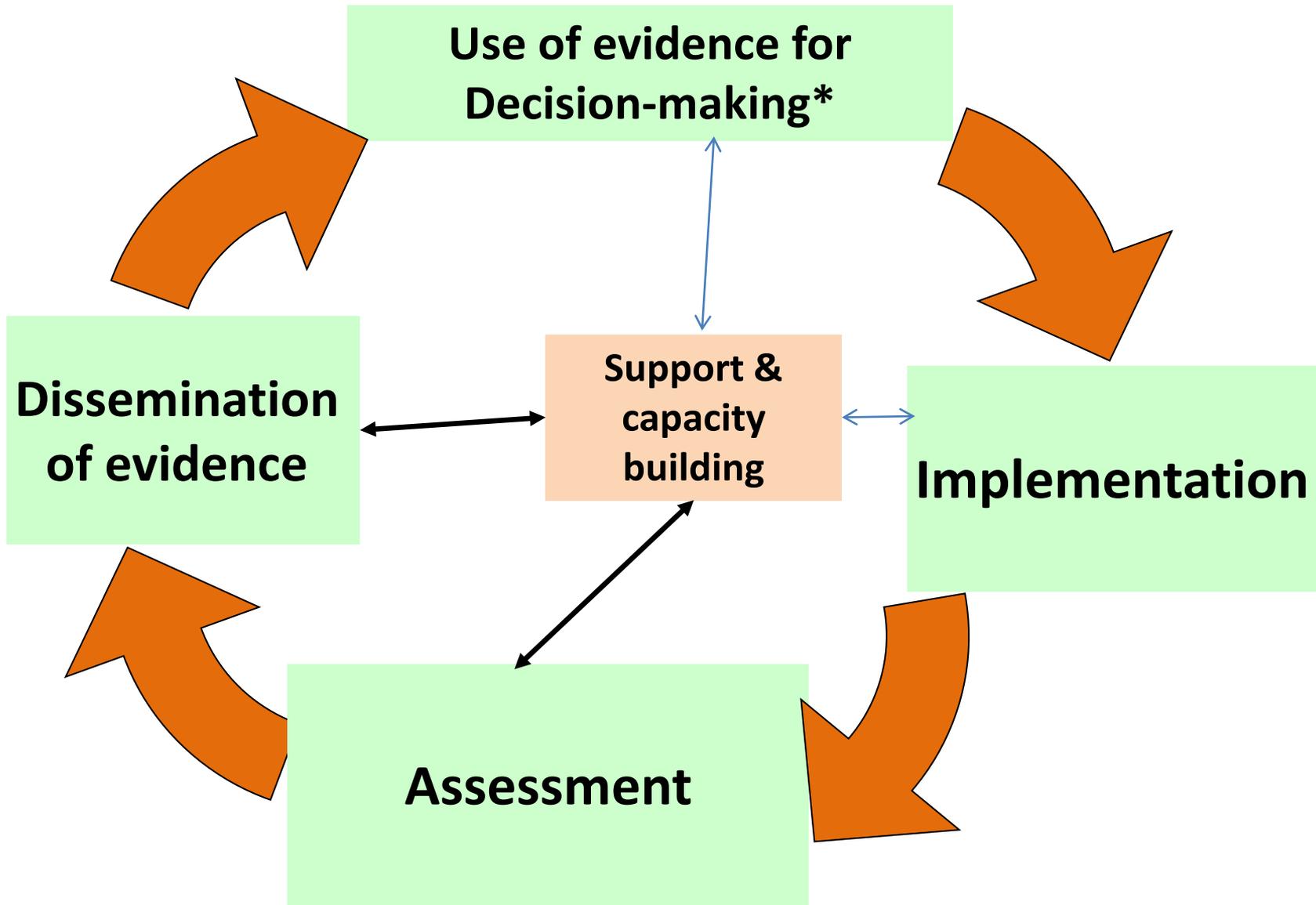
Purpose of Standards and Performance Level Descriptors (PLDs)

- To provide coherent information across grades and subjects
- To guide the development of the assessments
- To guide teacher's instructional efforts to ensure that students reach performance levels**s**

Impact of national assessments

- Uruguay is a model example of use of national assessments with proven positive impact about 5% improvement in performance in less than five years
- What did they do?:
 - Intensive communication of results
 - Support to schools and teachers (including weekend workshops and discussion groups)
 - Incentives to improving schools

Theory of Action and Policy Implications



Policy Implications

- Policy must spell out in unambiguous terms the purpose(s) of national assessments
- Strengthen policy implementation by building into the policy clear mechanisms for monitoring the implementation to avert possible obfuscation
- Report assessment evidence in ways that meet the needs of the end-users (no one-size-fits-all reporting)

Conclusion

- There's general agreement on the importance of assessment data to inform various decisions in education
- There's tendency to conveniently use assessment data for multiple purposes, particularly disproportionately more for accountability than for development
- South Africa has good opportunity to tread carefully, craft policy that is clear on the purpose(s) that national assessments must serve
- Monitor the feedback and utilisation of assessment data and track the impact on the system