

Using Student Assessments

What can we learn from Latin America?

Jeffrey M. Puryear
Inter-American Dialogue / PREAL
IEA/Hong Kong
October 9, 2007

I. The Status of Student Assessments

- **National testing systems have grown substantially**
 - 1980: non-existent
 - Today: almost universal

First National Assessments



Source: Sistemas de evaluación de aprendizajes en América Latina. Balances y desafíos. PREAL, 2006. 3

Diverse Institutional Arrangements

- Some are inside ministries; others are outside
 - Brazil, Colombia and Mexico are outside the ministry
 - Those outside ministries have been more stable, flexible and consistent, but no clear trend
- Some are autonomous, stable and prestigious; others are vulnerable to government interference
- Some are strong; others are weak

Subnational Tests

- Bogotá (Colombia)
- Paraná, São Paulo, Minas Gerais (Brazil)
- Buenos Aires Province and City (Argentina)
- Nueva León, Sonora and Aguas Calientes (Mexico)

Who and what are tested? (1)

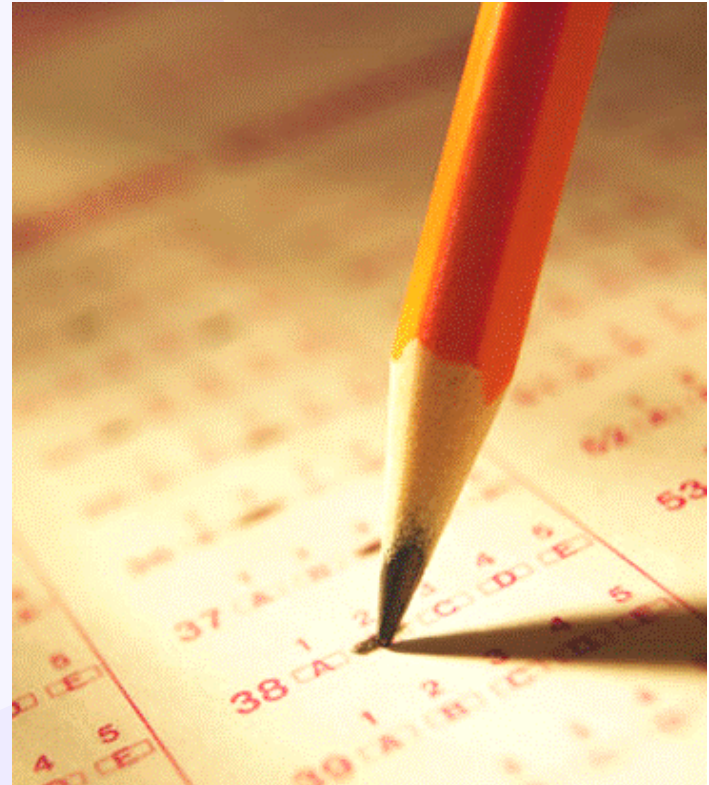
- Sample-based tests are the norm, but census-based tests are growing
- Started with math & language at primary level
- Today, many different levels and subjects are tested

Who and what are tested? (2)

- All use closed, multiple-choice tests; a few use open-ended questions
- Norm-referenced tests are being replaced by criterion-referenced tests
- Relationship of tests to curricular goals is often tenuous

Types of Testing

- Low-stakes tests are the norm
- High-stakes tests are appearing



Governments have not communicated test results well

- Top priority given to government officials
- General public usually not targeted
- Local education actors targeted only occasionally
- Parents usually ignored
- Press communicates more successfully to teachers, parents, and decision-makers

Participation in International Student Tests (1)

- Low but growing
- Global Tests
 - TIMSS, PIRLS, PISA are most common
 - Only a minority of countries have participated in a global achievement test
 - Participation is growing steadily

Participation in International Student Tests (2)

- Latin America-wide Tests
 - Latin American Laboratory for the Evaluation of Educational Quality (LLECE)
 - Established & managed by UNESCO/Latin America
 - Tests language, science & math at primary level
 - 1997 – 13 countries
 - 2006-2007 (SERCE-LLECE) – 20 countries

Participation in International Student Tests (3)

- Impact of tests has been weak
 - Have principally affected public opinion
 - Governments have not properly specified goals nor adequately communicated results
 - Low scores have sparked national concerns

Participation in regional & international tests

	FIMMS/ SIMMS/ TIMMS	IALS (1998)	UNESCO/ OREALC LLECE (1998)	IEA Civics Education (1999)	TIMMS-R (1999)	PISA (2000)	PIRLS (2001)	TIMMS (2003)	PISA (2003)	UNESCO/ OREALC SERCE (2006)	PISA (2006)	PIRLS (2006)	TIMMS (2007)
Argentina	X		X			X	X	X		X	X		
Belice							X						
Bolivia			X										
Brasil			X			X			X	X	X		
Chile		X	X	X	X	X		X		X	X		
Colombia	X		X	X			X			X	X		X
Costa Rica			X							X			
Cuba			X							X			
Ecuador			X***							X			
El Salvador										X			X
Guatemala										X			
Haití													
Honduras			X										X
Jamaica													
México	X		X			X			X	X	X		
Nicaragua										X			
Panamá										X			
Paraguay			X							X			
Perú			X			X				X			
República Dominicana	X		X							X			
Trinidad y Tobago												X	
Uruguay									X	X	X		
Venezuela			X										

Student assessments are not expensive

- Costs range from .02% to .33% of the national budget for the educational level tested (primary, secondary)
- Per-student costs vary greatly, depending on:
 - Number of subjects, grades & students tested
 - Intended use
 - Geography
 - Languages
 - Multiple-choice vs. open-ended questions
- Student assessments have proved to be one of the least expensive measures to reform education

II. Use of Student Assessments

1. Increase public awareness
2. Improve policy
3. Conduct research and program evaluations
4. Establish educational standards
5. Target schools or student groups for special support
6. Improve teaching
7. Identify individual children at risk
8. Promote accountability

1. Increase public awareness (1)

- Most common and effective use to date
- Tests have:
 - provided hard evidence that achievement is low and unequal
 - stimulated public concern that a serious problem exists
 - shifted policy focus from access to learning
- Media have played a key role

1. Increase public awareness (2)

- Beginning to alter the demand for education
 - Stimulating demand for more equitable distribution of opportunities to learn
 - Peru: 2003 declaration of a “state of emergency” in education
 - Uruguay: decision to target schools that serve less-advantaged students
 - Colombia: decision to establish modern education standards
 - Chile: Secondary students’ protest of 2006

2. Improve policy (1)

- Probably the most commonly expected use
- Few countries have used test data to improve policy
 - Chile a notable exception
 - Used test scores to shift reform efforts towards setting standards and providing support to teachers regarding pedagogical practice
 - Mexico, Uruguay, Bogotá, Sao Paulo

2. Improve policy (2)

- Variants of the U.S. No Child Left Behind program may be appearing (Chile, El Salvador, Brazil)
 - Setting learning standards
 - Testing to see whether children are meeting the standards
 - Developing interventions for schools that fail to meet standards

3. Conduct Research & Program Evaluation

- Availability of test scores has stimulated “determinants of learning” research
- Studies increased ten-fold between 1999 and 2005
- Governments initially reluctant to share data with research community

4. Establish national standards

- Largely unexpected
- Act of testing forced focus on what students are supposed to learn
- Assessment professionals have demanded clearer learning expectations
- Stimulated shift from norm-referenced tests to criterion-referenced tests
- Several countries now setting content & performance standards

5. Target schools or student groups for special support (1)

- Identify schools that most need assistance
 - Chile: Selected lowest performing 10% of schools for special support; Used test scores to establish “School Vulnerability Index”
 - Uruguay: Used test scores to determine which schools will have a longer school day
 - Bogotá: Identified schools for “Leveling for Success” program

5. Target for special support (2)

- Design school improvement plans that change the school's ethos & goals
 - Peru: Institutional Development Projects
 - Colombia: Institutional Education Projects & Improvement Plans
 - Uruguay: Educational Improvement Projects
 - Paraguay: Educational Innovation Projects
 - Chile: Educational Improvement Projects
- Easier to do with census-based testing
- This usage is growing

6. Improve teaching (1)

- Limited and experimental to date
- Several countries beginning to work with teachers to interpret test results
 - Chile delivers test results to individual schools with a guide to help analyze the results
 - Goal is to get teachers to reflect on their implications vis-a-vis causes and teaching practices
 - Most teachers value them, but around 40% do not find them useful
 - Peru beginning to use tests to provide teachers with suggestions of what kinds of learning they must promote

6. Improve teaching (2)

- Tests have an impact on teachers' perceptions:
 - What is tested will be taught
 - Criterion-referenced tests help teachers visualize what their students must be able to do
- However, no evidence that test results are being used to design training courses for teachers

7. Identify Individual Children at Risk

- Only a few initial experiments
 - El Salvador
 - Chile
- Requires relatively frequent census-based tests

8. Promote Accountability (1)

- Works best with census-based tests
- High-stakes tests are not common, except for Costa Rica, Dominican Republic and El Salvador
- Identifying schools for rewards or sanctions
 - Chile publishes all schools' average test scores (publicity matters)
 - El Salvador publishes scores of the highest-scoring schools
 - Bogotá, Colombia uses test scores to select schools for its "Excellence Award"

8. Promote Accountability (2)

- Identifying teachers for rewards or sanctions
 - New, limited and highly controversial
 - Chile gives salary bonuses to all teachers in high-performing schools (test scores + management indicators, adjusted for SES)
 - Mexico gives salary bonuses to individual teachers based on test scores plus other variables (questionable impact)

8. Promote Accountability (3)

- Setting national goals for student achievement
 - Still incipient
 - Chile & Bogotá have set minimum improvement targets for all schools
 - Advantages
 - Easy for public (and media) to understand, mobilizes public opinion
 - Schools compete against past performance rather than other schools

8. Promote Accountability (4)

- **Facilitating school choice**
 - Almost non-existent
 - Chile is the major exception
 - Ministry distributes school-by-school test results for each region, including
 - average scores in each subject area,
 - changes in scores since previous testing round, and
 - comparisons with schools serving similar socioeconomic groups
 - Special reports for parents since 2003
 - Little evidence that parents use test scores

8. Promote Accountability (5)

- Tests seldom used to promote accountability because
 - a new tool, not well-understood
 - often resisted by teachers and administrators
 - doubts about technical validity of tests

Bottom Line: Usage is hard to predict

- Most governments established testing systems without a clear idea of how they would use them
- Availability of data has prompted diverse (and often unexpected) uses
- Governments should think more carefully in advance how they might use test results

III. Emerging Lessons

- Student assessments have great potential that is far from being realized
- Usage is hard to predict
- Achievement tests help identify problems
- The first & most likely impact is on public opinion
- Test scores don't automatically influence policy
- Test scores stimulate debate on standards
- Governments don't disseminate results well
- Test scores may promote accountability, but not right away
- Achievement tests are not expensive

Keep in Mind:

- Test results are probably being used more than we suspect
- Despite absence of clear objectives, once information is available, people tend to find uses for it



www.preal.org