Setting an Accountability System in Israel

Michal Beller
Director-General of RAMA
The National Authority for Measurement and Evaluation in Education, Israel

http://rama.education.gov.il
The National Educational Plan

Because every child deserves more

"If you want to reform the world - you must reform education"

Janusz Korczak

September 2003 – December 2004
Principles of the Plan

- Decentralization: Accountability, Transparency
- Strengthening early stages of education
- Strengthening the schools: a warm, protective home
- Goal-oriented, Results-based education
- Strengthening public education
- High-quality pedagogical continuum
- Improving the teaching profession and its status
- Responsibility for every child
- Boosting achievement
- Reducing gaps
- Pooling of resources: streamlining, and realistic budgeting
An independent statutory unit - the National Authority for Measurement and Evaluation – RAMA - shall be established as the entity that leads and provides professional guidance to the education system with respect to measurement and evaluation.

RAMA will conduct periodic evaluations of the education system and evaluations in schools, and will publish its findings in an annual report submitted to the National Council for Education.

From the Dovrat National Task Force in Education (2004)
Introduction to RAMA Accountability - General Framework

Assessment for Learning - Large Scale Assessment in Israel

Updated System

Concluding Remarks

Negative Effects

National Authority for Measurement and Evaluation in Education

Assessment for Learning
Introduction

RAMA

Accountability - General Framework

Does it Work

Negative Effects

Large Scale Assessment in Israel

Updated System

Concluding Remarks

Assessment for Learning

Alignment with the goals of the educational system

"Measurement"

"Learning"

Ongoing feedback from the field

Summative

Formative

Internal

External

"Learning for Measurement"

RAMA
Accountability: A General Framework
Accountability Should Provide Support for an Ongoing Process

Baker, 2005
Ultimately, accountability is not only about measuring student learning but actually improving it.

Consequently, genuine accountability involves supporting changes in teaching and schooling that can heighten the probability that students meet standards.

There are at least three major areas where attention is needed:

- Ensuring that teachers have the knowledge and skills they need to teach to the standards.
- Providing school structures that support high quality teaching and learning.
- Creating processes for school assessment that can evaluate students’ opportunities to learn and can leverage continuous change and improvement.

Darling Hammond, 2004
Does External Accountability Affect Student Outcomes?
The Jury is Out

- Henry Braun (2004) conducted a study in which he used four different modes of analysis to evaluate the data on the connections between statewide high-stakes testing and student achievement.

- He concluded that the decisions that researchers made about methods of analysis largely determined which kinds of findings they reported:
  - analyzed in some ways, the evidence showed positive effects for high-stakes testing;
  - analyzed in other ways, there was no discernible effect.

Shulman, 2006
What Works and When?

- Accountability systems work to the degree that they engage the knowledge, skill, and commitment of people who work in schools.

- The success of accountability policy depends on the development of practices of improvement—explicit strategies for developing and deploying knowledge and skill in classrooms and schools.

- The politics of accountability tend to lead to an underinvestment in knowledge and skill, and an overinvestment in testing and regulatory control.

- Correcting this distortion requires changing the relationship between policy and practice, particularly around the definition and development of leadership.

Elmore, 2006
Can We Measure Learning When Stakes are High?

- Information from test-based accountability (TBA) systems is likely to be more valid under some conditions than others.

- But scores on high-stakes tests will always provide an incomplete indicator of learning.

- Developers of TBA systems must create conditions that will promote valid information while informing users of the limitations of this information.

- Recent research suggests that TBA has some potential to promote positive outcomes, but efforts must be taken to reduce likelihood of negative consequences.

Hamilton, 2005
Unintended Negative Effects of External Testing

- Between-Subject Reallocation of time
- The Impact of Setting a Specific Target
- “Most Children are Left Behind…”
- Teaching to the Test; Test Inflation
- Narrowing the Curriculum
- Assessment is Only a Sample of the Curriculum
Large Scale Assessment in Israel
Large Scale Assessments in Israel

- Matriculation Exams (‘Bagrut”) – end of high school

- Meitzav (Mother Language, Math, Science, English) – 5th and 8th grades (+2nd grade in Language)

- International Assessments:
  - TIMSS - Math and Science Study
  - PIRLS - Reading Assessment
    • (2001, 2006) – 41 countries
  - SITES - Information Technology
    • (1997, 2006)
  - PISA - Reading, Mathematical & Scientific Literacy
    • (2000/2, 2006) – 58 countries
Meitzav
National Assessments
Meitzav - Growth and Effectiveness Measures for Schools (GEM) - introduced by the Ministry of Education in 2002/3

- A set of school level indicators, with a goal to enable principals to base their decisions on valid evidence
- Administered every second year to all 5th and 8th (+ 2nd grade for Mother Language)

The Meitzav School Report includes information on:

- **Pedagogical Environment** in the School (based on questionnaires and interviews)
- **School Climate and Work Environment** (based on questionnaires and interviews)
- **Student Achievements** –
  - Based on results from standardized national assessments, administered in both Hebrew and Arabic
  - Math, English, Mother Language (Hebrew/Arabic), and Science & Technology
Alignment

Teaching & Learning

Content & Performance Standards

Curriculum

Educational Environment

Achievement Tests

School Climate

Introduction

RAMA

Accountability - General Framework

Does it Work

Negative Effects

Large Scale Assessment in Israel

Updated System

Concluding Remarks

© 2007 # 19
School Climate & Pedagogical Environment
Introduction to RAMA

Accountability - General Framework

Does it Work

Large Scale Assessment in Israel

Updated System

Concluding Remarks

Negative Effects

Questionnaires & Interviews

Students

Teachers

Principals

Educational Environment

School Climate
Updating the Israeli Evaluation System
Untended Negative Effects of Meitzav

- Between-subjects reallocation of time
- Teaching to the test
  - Narrowing of the curriculum
  - Excessive test preparation
  - Score inflation
- Gaming the system (incl. cheating)
- Excluding weak students
The main **goals** of the recent update:

- Implement a culture of **“assessment for learning”**
- Mitigate the **threats** of external exams (including gaming the system)
- Effective integration of internal and external evaluation
- Decentralization of the evaluation process along with the use of centrally designed rigorous tools
- Empowerment of teachers and principals
- Professional design of assessments and scoring (including equating, alignment to standards and more)
Use of Multiple Indicators

- Internal: Formative Assessments, Assignment Data Base, “Off-the-shelf” Assessments
- Internal admin. of external tests
- External: Matriculation exams, Meitzav

Schools & pupils are being examined

Sample-based Assessments (like NAEP & International exams)
## Cycles of Internal & External Meitzav

<table>
<thead>
<tr>
<th>Cluster of Schools</th>
<th>Subject</th>
<th>2006/7</th>
<th>2007/8</th>
<th>2008/9</th>
<th>2009/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Science &amp; Technology</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Math</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Mother Language</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
</tr>
<tr>
<td>B</td>
<td>Science &amp; Technology</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Math</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td></td>
<td>Mother Language</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td>C</td>
<td>Science &amp; Technology</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Math</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Mother Language</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td>D</td>
<td>Science &amp; Technology</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>Internal</td>
<td>Internal</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td></td>
<td>Math</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Mother Language</td>
<td>Internal</td>
<td>External</td>
<td>Internal</td>
<td>Internal</td>
</tr>
</tbody>
</table>
Goals of the Internal Meitzav

- **Reduce the frequency of external assessments** (from every 2 years to every 4 years), and introduce internal assessments in between

- **Provide principals with a way to annually ensure** that teaching is aligned with expectations and standards for each subject matter

- **Annual follow up of changes in achievement and school climate relative to national norms**

- **Enable teachers to internally use professional assessment tools** to plan and monitor learning

- **Provide feedback at the student level** (hence increase the motivation of students to respond)

- **Implement a culture of “assessment for learning”** and focus attention on aspects not measured by standardized tests
Concluding Remarks
It Is All About Teacher Quality…

- The most likely way to improve student performance is to improve the quality of teachers

- Rivkin, Hanushek and Kaine (2005) estimate that the differences in annual achievement growth between an average and a good teacher are large.

- Within one academic year, a good teacher can move a typical student up at least four percentiles in the overall distribution (equal to a change of 0.12 standard deviations of student achievement).

- The difficulty, is that hiring good teachers is not easily done. Teaching ability is not closely related to training or experience. Moreover, common salary systems do not target particularly high quality teachers

- From a policy viewpoint the primary objective should be improving the overall quality of the teaching force

Hanushek, 2005
Introduction
RAMA
Accountability - General Framework
Does it Work
Negative Effects
Large Scale Assessment in Israel
Updated System
Concluding Remarks

Trust and Shared Responsibility

Wider Community

Active Partners

External Stakeholders

Students
Parents
Teachers
Politicians
Administrators
Local Education Units

Trust and Shared Responsibility
Thank You

mbeller.rama@education.gov.il

Comments ? Questions?