The role of educational monitoring, assessment, and standards in school reform

New policy strategies in the German speaking countries

Eckhard Klieme

Educational Policy and Reform in a Global Age
International Seminar, December 11-12, 2004
Center for Research of Core Academic Competencies
Graduate School of Education, The University of Tokyo
1. The „PISA shock“ and beyond

2. The new policy strategy: standard-based quality management

3. Issue #1: Models of competence - The core of educational standards

4. Issue #2: School accountability - Multiple perspectives on evaluation
A typology of educational cultures
(Helmut Fend)

Support structures for teachers and schools

- Good will and policy
- Deprivation
- Control of Standards
- Integrative quality management
- External control
Educational Traditions in Germany

• Normative basis: „Bildung“ as process of personal development + cultural participation

• Teacher as autonomous professional, introducing students into the core ideas of his subject

• 3 tracks in lower secondary schools: different mixture of „Bildung“ and vocational training

• No Market; equality assumed

• No centralized test system
Thesis on inter-national comparison

• Gross level in educational performance is determined by general cultural and socio-economic factors.

• Structure of the educational system determines (a) between-school-variation and (b) SES impact.

• Educational and instructional culture determines profile of outcomes (relative strengths and weaknesses).

→ Comparative research has to study relations and profiles rather than levels.
### Relative Strengths and Weaknesses in Mathematics

Identified by differential item functioning (DIF) parameters in TIMSS and PISA, correlated with demands.

<table>
<thead>
<tr>
<th>Country</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>Geometry</td>
<td>Extra-mathematical contexts; graphs</td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>-</td>
<td>Technical skills</td>
</tr>
<tr>
<td>France</td>
<td>Technical skills</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>„realistic mathematics“</td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>Combining representational formats</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>
PISA vs. TIMSS

different concepts of mathematical competence

Source: Expert ratings in PISA 2000/Germany
From PISA 2000 to PISA 2003

Rank orderings of countries are no valid indicator of changes. Equating of test scores from PISA 2000 and 2003 shows for Germany:

+ science
+ „functions“ / = „space“ in mathematics
= reading

Interpretation:
Reforms stimulated by TIMSS (and later also by PISA 2000) show first effects.
School climate (PISA 2000)

- Pressure to perform
- Support
- Quality of relation

Countries:
- United States
- New Zealand
- United Kingdom
- Australia
- Canada
- Sweden
- Finland
- Norway
- Japan
- Luxemburg
- Belgium
- Netherlands
- France
- Switzerland
- Austria
- Germany
- Poland
- Hungary
- Russ. Federation
- Korea
### Impact of school factors in PISA 2000

<table>
<thead>
<tr>
<th>Country</th>
<th>N of schools</th>
<th>SES</th>
<th>Disciplinary problems</th>
<th>Teacher support</th>
<th>Quality of relation</th>
<th>Achievement press</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>152</td>
<td>.70 ***</td>
<td>-.09</td>
<td>-.25 ***</td>
<td>.11 *</td>
<td>.00</td>
<td>.534</td>
</tr>
<tr>
<td>Finland</td>
<td>155</td>
<td>.35 ***</td>
<td>-.20 **</td>
<td>.08</td>
<td>-.25 **</td>
<td>-.24 **</td>
<td>.234</td>
</tr>
<tr>
<td>Japan</td>
<td>135</td>
<td>.41 ***</td>
<td>-.36 **</td>
<td>-.13</td>
<td>.18</td>
<td>.06</td>
<td>.372</td>
</tr>
<tr>
<td>Switzerland</td>
<td>279</td>
<td>.43 ***</td>
<td>-.12 *</td>
<td>-.26 ***</td>
<td>.12</td>
<td>-.16 **</td>
<td>.377</td>
</tr>
<tr>
<td>Germany</td>
<td>1.473</td>
<td>.71 ***</td>
<td>-.18 ***</td>
<td>-.21 ***</td>
<td>.06 **</td>
<td>-.02</td>
<td>.702</td>
</tr>
<tr>
<td>Austria</td>
<td>212</td>
<td>.67 ***</td>
<td>-.03</td>
<td>.02</td>
<td>-.08</td>
<td>.03</td>
<td>.450</td>
</tr>
<tr>
<td>Korea</td>
<td>146</td>
<td>.18 ***</td>
<td>-.27 ***</td>
<td>-.24 **</td>
<td>.10</td>
<td>.55 ***</td>
<td>.548</td>
</tr>
</tbody>
</table>
Summary of main issues in PISA 2000/Germany:

• low achievement and high SES impact
• unsatisfactory profile of achievement and school climate
• large differences between German Länder
• large variations in grading requirements

Linked by empirical arguments from PISA and SER

Policy reactions in Germany:

• new pedagogical initiatives:
  foster reading, deal with heterogeneity
• all-day schooling
• professionalization, esp. for elementary teachers
• national standards, system monitoring, and evaluation
Standards for Student Competencies (Klieme et al. 2003)

• are aligned to general educational aims
• describe competencies and their levels by constructing models of competence
• operationalize levels of competence by tasks and test items
• specify expected levels for certain grades
• serve as criteria for school accountability
The economics and politics of standard-based educational reform
(Klieme et al. 2003)

Complex models of competence: new orientation for teacher

Complex assessment (including cross-curricular tasks, student portfolios etc.)

System monitoring – School evaluation – student diagnostics aligned, but not identical (no high stakes testing)

Grading and certification done by the teacher

School development: Focus on professional development

School evaluation: Multiple perspectives and methods
Issue # 1: Models of competence –
The core of educational standards

Chomsky/Piaget: competence = generic mental structure

Functional concept of competence (e.g., Bandura 1990):
using knowledge and abilities in ill-defined, unpredictable, stress-inducing situations
Issue # 1: Models of competence – The core of educational standards

Vocational education: The search for generic skills, which support employability & life long learning

„Schlüsselqualifikationen“ (Mertens)
Key competencies (NL, EU)
Core competencies (UK, USA: SCAN)
Life skills (OECD, DeSeCo)

• Basic competencies (reading, numeracy, ICT)
• Transferable competencies (problem solving)
• Competencies which support transfer (self regulation)
Issue # 1: Models of competence –
The core of educational standards

Weinert(2001, S. 27f.): Definition of „competencies“

skills and abilities required to to solve certain kinds of problems,

connected to motivational, volitional, and social abilities
and intentions required to apply problem solutions within a
wider range of situations
Instruments:
Analytical Problem solving (PISA 2003)

<table>
<thead>
<tr>
<th></th>
<th>Angaz</th>
<th>Kado</th>
<th>Lapat</th>
<th>Megal</th>
<th>Nuben</th>
<th>Piras</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>550</td>
<td>300</td>
<td>850</td>
<td>550</td>
<td>450</td>
<td>250</td>
</tr>
</tbody>
</table>

Note: The values represent scores or measurements in the context of the instruments used for assessing analytical problem-solving skills in PISA 2003.
Instruments:
Analytical Problem solving (PISA 2003)

3 levels of problem solving competence in PISA

Comprehend a system or device, identify parts/variables and constraints, induce features
(step-by-step reasoning, mainly small discrete systems)

Make a decision with respect to a well-defined set of criteria or analyse a system, combine representations
(constraint-based reasoning, continuous systems)

Design a system as a whole or make a diagnose / propose a solution, reflect and communicate
(causal or combinatorial reasoning, complex qualitative systems)
Issue #2: School accountability
Multiple perspectives on evaluation

- Surveys
- Standard-based assessment
- Audit, Peer Review

- Interactive
- Reflexion

- Summative
- Administrative / External
- Formative
- Participative / Internal
Knowledge Base (based on Scheerens & Bosker 1997): School effectiveness and instructional quality

**Context**
School structure, Curriculum, pedagogical traditions and orientations, teacher education, budgeting and regulation, socio-economical and cultural context

<table>
<thead>
<tr>
<th>INPUT</th>
<th>PROCESS</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ressources: Teacher-student-rate, student population, parent commitment</td>
<td>School leadership, teacher cooperation, cohesion, norms and values, school climate, evaluation</td>
<td>School level</td>
</tr>
<tr>
<td>Students per class, teacher competencies</td>
<td>Instructional quality: -- classroom management -- student orientation -- cognitive activation</td>
<td>Classroom level</td>
</tr>
<tr>
<td>SES, social and cultural capital, family support, gender, language and migration background, general intellectual ability, pre-knowledge</td>
<td>Motivation and interest, self concept, learning strategies, self regulation, time invested</td>
<td>Individual level</td>
</tr>
</tbody>
</table>
1. The „PISA shock“ and beyond

2. A quest for understanding the inter-national variation

3. The new policy strategy: standard-based quality management

4. Issue #1: Models of competence - The core of educational standards

5. Issue #2: School accountability - Multiple perspectives on evaluation

Needed: Empirical Research on effects of strategic reforms
Donald Campbell 1969
(American Psychologist, 24, 409-429)

Reforms as Experiments

„It is one of the most characteristic aspects of the present situation that specific reforms are advocated as though they were certain to be successful.“

„We must help create a political climate that demands more rigorous and less self-deceptive reality testing.“
(I) founded in 1951 with U.S. support to build a democratic school system by fostering empirically based pedagogics

(II) outstanding for comparative educational research, esp. in qualitative and historical paradigms

(III) since 1999 restructured as Centre for
a) Information on educational issues to the public
b) Evaluation and system monitoring (incl. international comparisons)
c) Research on quality of education on the system, school, and instructional level