Quality of (Schools and) Teaching: What Can we Learn From International Studies?

Eckhard Klieme

Deutsches Institut für Internationale Pädagogische Forschung

DIPF

Educational Research
and Educational Information

IEA Research Conference
Singapore, June 28, 2013

Outline

Challenges in international LSA

- Understanding the (causal ?) impact of professional practices and institutional contexts

- Distinguishing between classroom and school level
  ($\rightarrow$ Design !)

- Understanding the impact of system and culture

- Understanding the potential and the limits of cross-sectional data

- Measuring and explaining change on different levels
  ($\rightarrow$ Design!)

- Linking to strong theories of school /teaching effectiveness

Outline

1. Current challenges of international large scale assessment
2. Teaching in IEA studies: From SIMS to TIMSS Video
3. Theoretical approaches and findings on „teaching quality" -> OTL (content coverage), effective methods, process quality
4. Explorations of Teaching practices, quality, and effectiveness based on TIMSS 2003, grade 8 Mathematics data →

   a) three-level HLM of cognitive and non-cognitive effects - introducing science achievement as a control variable

   b) understanding cultural profiles in math teaching
5. Combining TIMSS/8 and PISA, understanding the difference
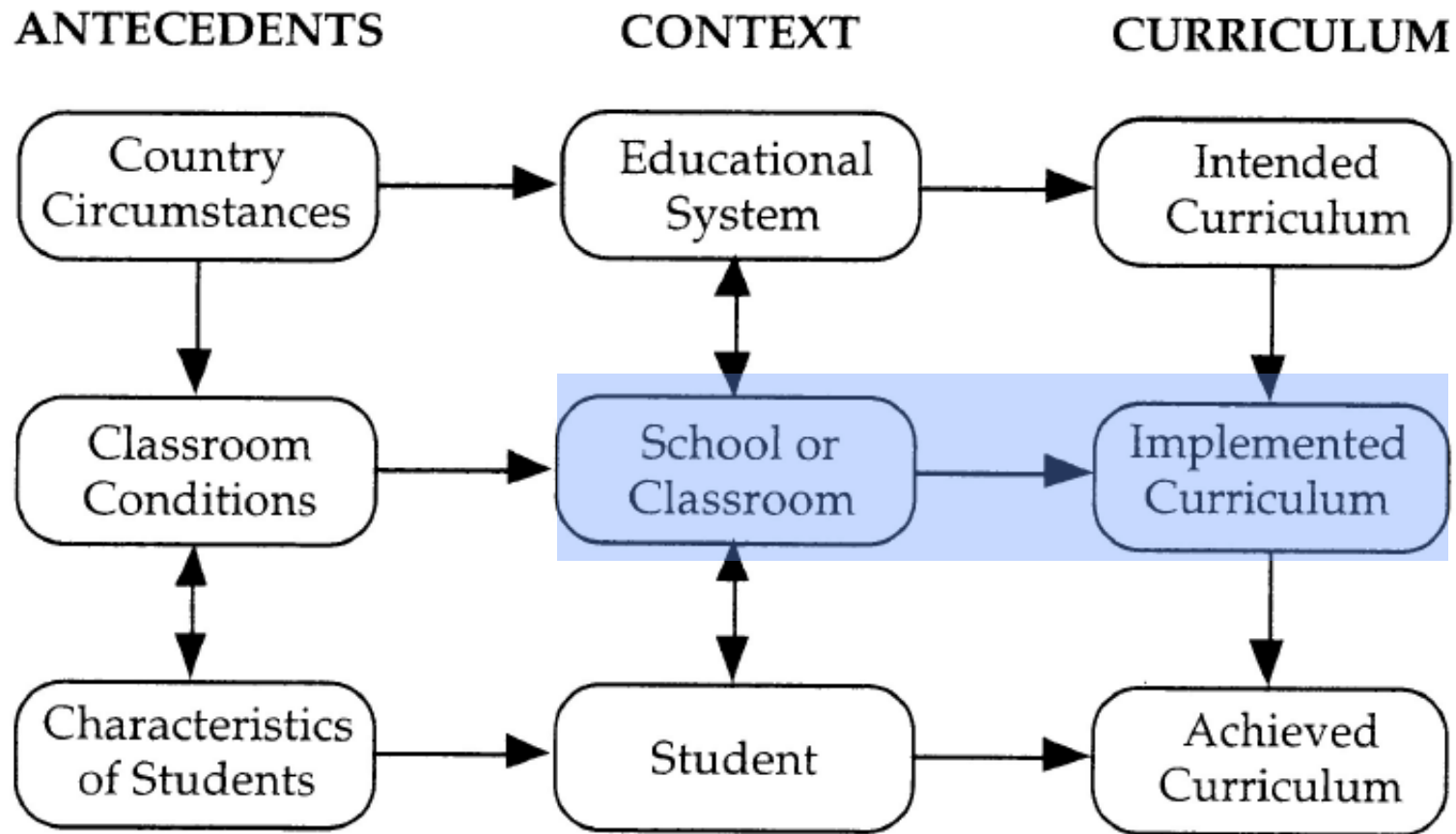6. Some suggestions for research stratgies in IEA

Conceptual Background (IEA, 1971)

ANTECEDENTS	CONTEXT	CURRICULUM
Country Circumstances	Educational System	Intended Curriculum
Classroom Conditions	School or Classroom	Implemented Curriculum
Characteristics of Students	Student	Achieved Curriculum

The „CIPO-Model"

CONTEXT
School structure, Curriculum, pedagogical traditions and orientations,
teacher education, budgeting and regulation, socio-economic and cultural context

INPUT	PROCESS	OUTPUT
Teacher-student-rate, qualification of teaching staff, student population, parent commitment	Quantity of instruction, School curriculum, leadership, teacher cooperation and collaboration, professional development, cohesion, school culture (norms and values), school climate, internal and external evaluation	**School level**
Students per class, teacher competencies	Instructional quality: - opportunity to learn - clear, well-structured classroom management, - supportive, student-oriented classroom climate, - cognitive activation with challenging content	**Classroom level**
SES, social and cultural capital, family support, gender, language and migration background, general intellectual ability, pre-knowledge	time invested, self regulation, motivation and interest, self concept, learning strategies	**Individual level**

1970-71/1976

The IEA
Six Subject Survey:
An Empirical Study
of Education in
Twenty-One Countries

David A. Walker

With Contributions by
C. Arnold Anderson and
Richard M. Wolf

With a Foreword by
Torsten Husén

Almqvist & Wiksell International
Stockholm, Sweden

A Halsted Press Book
John Wiley & Sons
New York – London – Sydney – Toronto

„Conducting classes in English was moderately associated with student achievement at the upper secondary school level and weakly associated with achievement at the 14-year-old level."

1970-71/1976   1981-1983/1989

The IEA
Six Subject Survey:
An Empirical Study
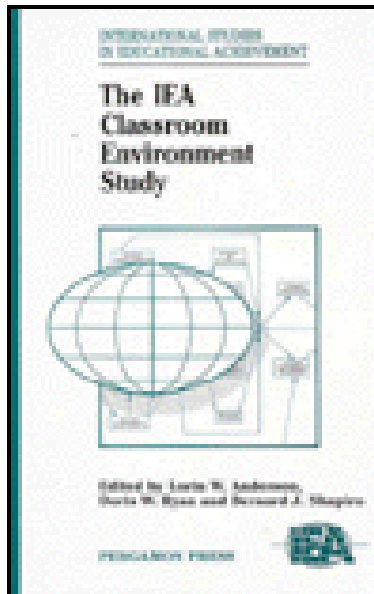of Education in
Twenty-One Countries

David A. Walker

With Contributions by
C. Arnold Anderson and
Richard M. Wolf

and a Foreword by
Torsten Husén

Almqvist & Wiksell International
Stockholm, Sweden

A Halsted Press Book
John Wiley & Sons
New York – London – Sydney – Toronto

INTERNATIONAL STUDIES
IN EDUCATIONAL ACHIEVEMENT

The IEA
Classroom
Environment
Study

Edited by Lorin W. Anderson,
Doris W. Ryan and Bernard J. Shapiro

PERGAMON PRESS   IEA

"Teacher behaviors were more consistently associated with students' academic engagement than with their final achievement. Thus, what teachers did in the classroom appeared to be more highly related to what students did than to what they learned."

IEA Classroom Environment Study
Anderson et al. 1989
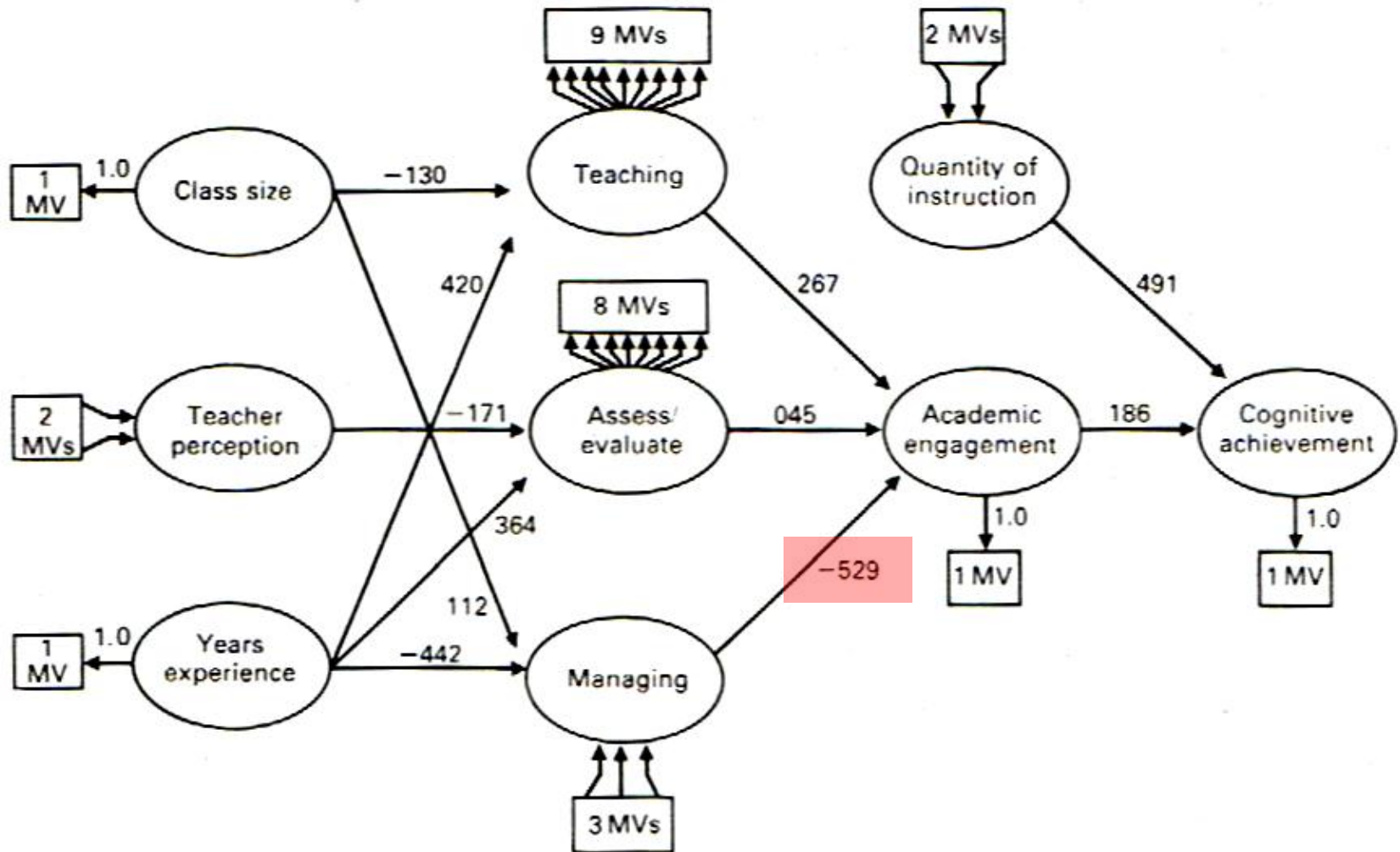
1970-71/1976   1981-1983/1989   1980-82/1993

The IEA
Six Subject Survey:
An Empirical Study
of Education in
Twenty-One Countries

David A. Walker

With Contributions by
C. Arnold Anderson and
Richard M. Wolf

And a Foreword by
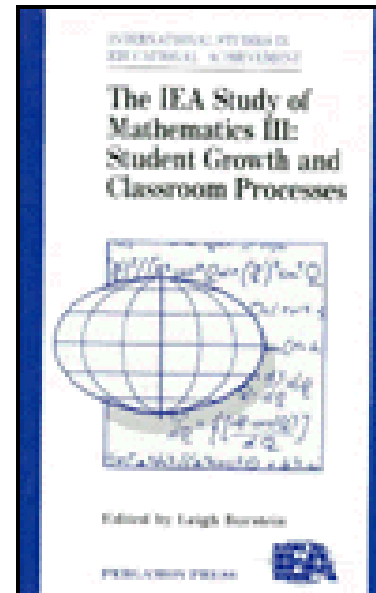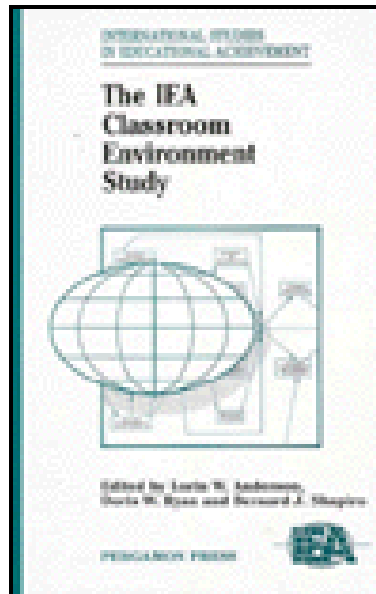Torsten Husén

Almqvist & Wiksell International
Stockholm, Sweden

A Halsted Press Book
John Wiley & Sons
New York – London – Sydney – Toronto

INTERNATIONAL STUDIES
IN EDUCATIONAL ACHIEVEMENT

The IEA
Classroom
Environment
Study

Edited by Lorin W. Anderson,
Doris W. Ryan and Bernard J. Shapiro

PERGAMON PRESS IEA

INTERNATIONAL STUDIES IN
EDUCATIONAL ACHIEVEMENT

The IEA Study of
Mathematics III:
Student Growth and
Classroom Processes

Edited by Leigh Burstein

PERGAMON PRESS IEA

„The picture changes dramatically, however, once controls for entering
performance are introduced …. Almost all of the … differences disappear" (p
275)
„The only classroom or school variable to be significantly related to
achievement growth (controlling for other student and schooling variables) in
more than one system was opportunity to learn the content represented on
the test (OTL). Even for OTL the results are spotty and inconsistent."(320)

Opportunity to learn
(Schmidt & McKnight 1995; Schmidt & Maier 2009)

- Content Coverage

- Content Exposure Variables:

  considering time and depth of teaching

- Content Emphasis Variables:

  e.g., lower vs. higher order skills

1970-71/1976   1981-1983/1989   1980-82/1993   1995/1999
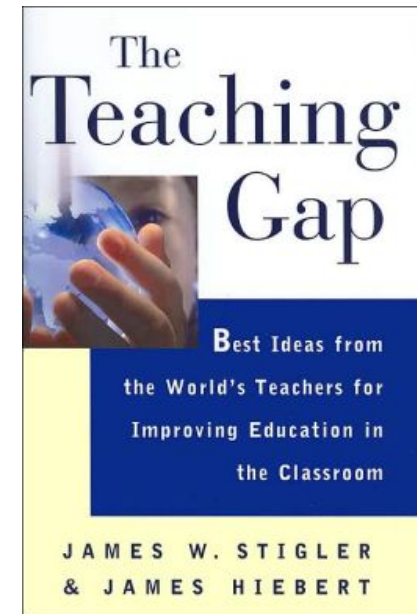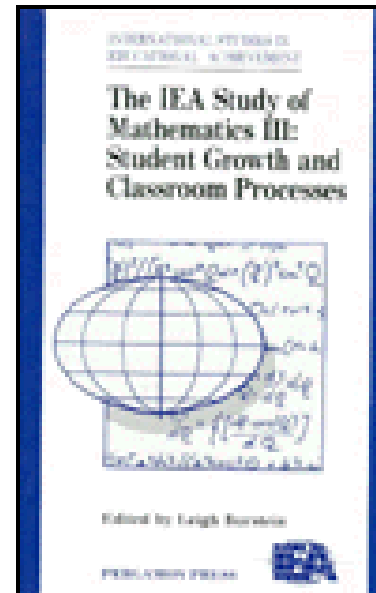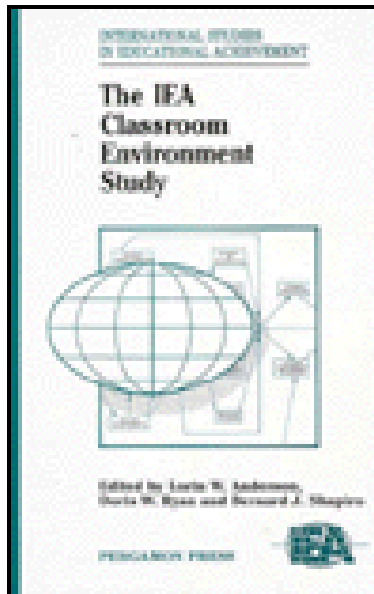
Scripts = generalized knowledge about an event that resides in the heads of participants.

These scripts guide behavior and also tell participants what to expect.

…learned through observation and participation.

Within a culture, these scrips are widely shared, and therefore they are hard to see.   (p. 85)

Teaching as „cultural activity" !   (Stigler)

**Percent of teaching time used for
challenging tasks vs. practicing (TIMSS Video Study)**
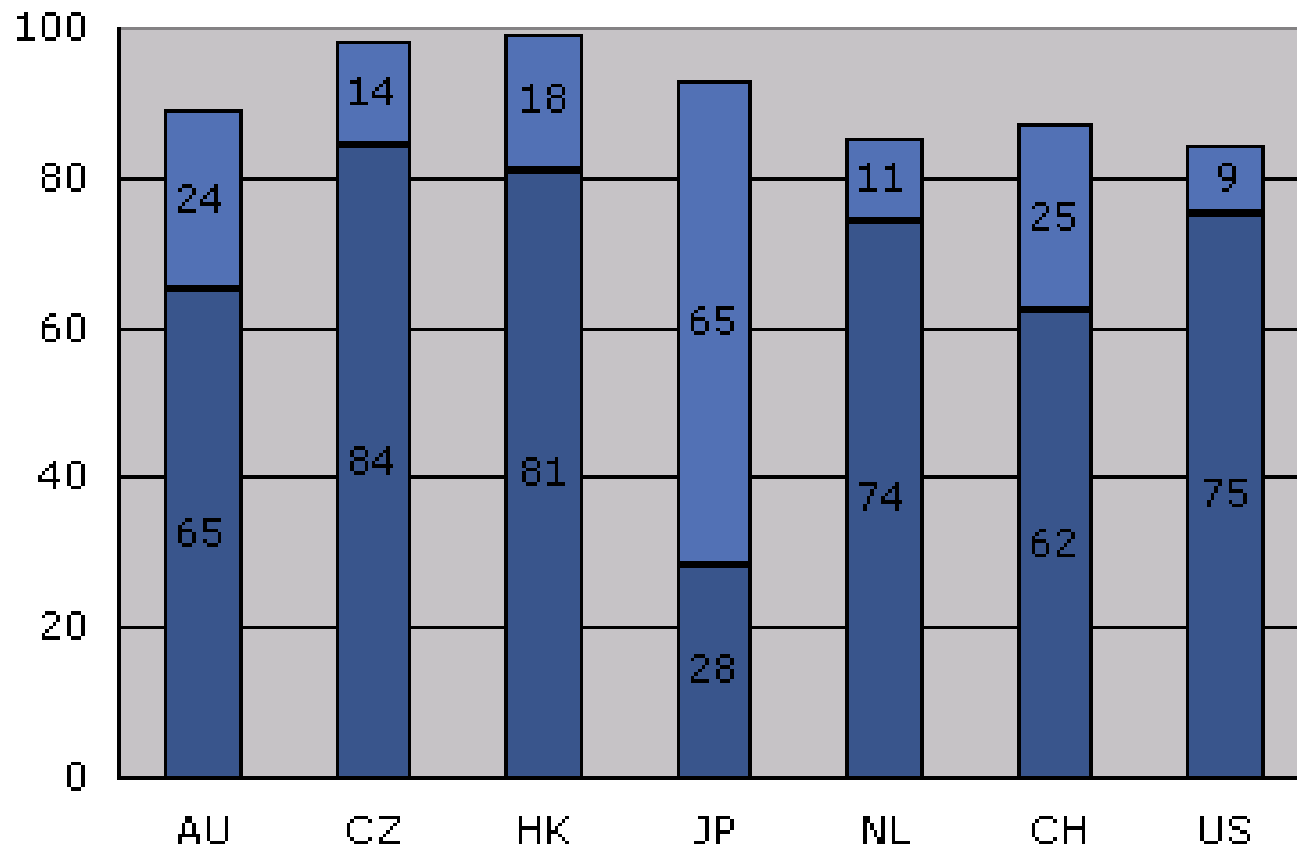
Teaching as „cultural activity" ? (Baker/Le Tendre)

For core instructional behaviors, most of the variation occurred within nations, not between nations.

Baker; D. & Le Tendre, G. (2005): The universal math teacher, p. 114

2nd order factors of classroom practice
based on high-inference video-ratings (Clausen,Klieme & Baumert 2002

(TIMSS-Video 1994 Germany: national sample, 100 + 86 lessons)

Structure and Classroom Management	Supportive climate	Cognitive Activation
Effective treatment of interruptions „teacher intervenes immediately, before disturbance may evolve" Clarity of rules Interruptions (-) Waste of time (-) Monitoring Time on task Teacher Unreliability (-) **Clarity and structuredness of the Instruction**	Social orientation: „teacher takes care of his students' problems" Teachers diagnostic competence with regard to social behavior **Individual reference norm in evaluation** Rate of interaction (-) Pressure on students (-)	Teacher's ability to motivate students: „can present even abstract content in an interesting manner " **Errors as opportunities** **Demanding tasks** Practicing by repetition (-)

Profiles for regions within Germany
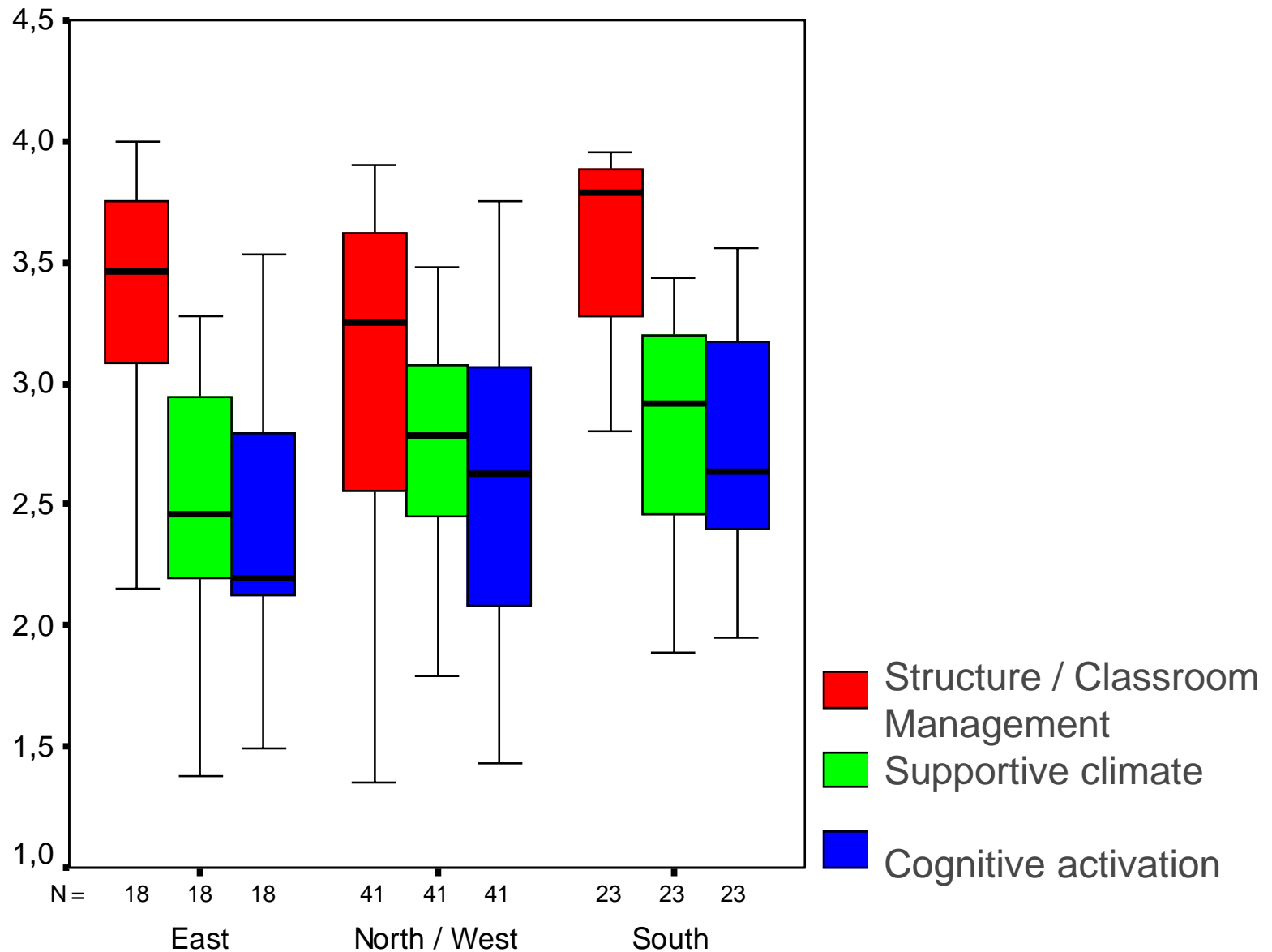
Correlations with types of utterances

Structure and classroom management	Supportive climate	Cognitive activation
More public talk More teacher utterances *High proportion of teacher utterances:* Related to mathematical content Questions, especially about facts Reactions to student answers Positive evaluations	Less public talk Fewer teacher utterances *High proportion of teacher utterances* questions: describe/explain personal ideas	More public talk *High proportion of teacher utterances* related to mathematical content questions: decisions, describe/explain

Supportive Climate predicts motivational development

Change in math interest over 1 year

Supportice Climate
(one selected lesson)

Hypothesis
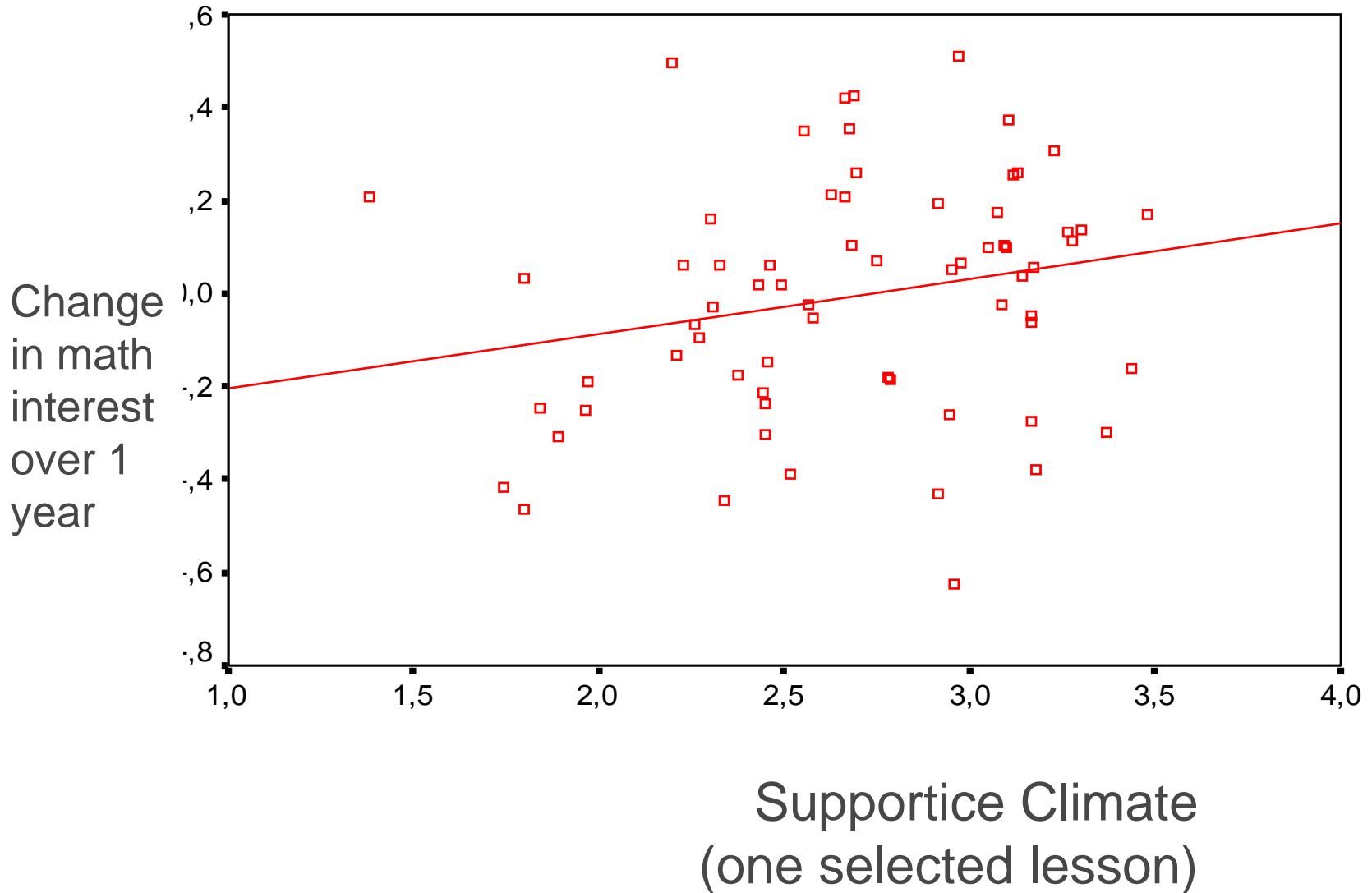
- Structure/Classroom Management,

- Support and

- Challenge/Cognitive Activation

are <u>basic featues of teaching</u> that encompass the

<u>quality of learning opportunities</u> (as opposed to subject matter)

in a non-redundant and non-reducible way.

These dimensions describe **how** the „implemented curriculum" is

implemented in classroom interaction.

(see: Raudenbush: „enacted instructional regimes")

Outline

1. Current challenges of international large scale assessment

2. Teaching in IEA studies: From SIMS to TIMSS Video

3. Theoretical approaches and findings on „teaching quality" -> OTL (content coverage), effective methods, process quality

4. Explorations of Teaching practices, quality, and effectiveness based on TIMSS 2003, grade 8 Mathematics data →

   a) three-level HLM of cognitive and non-cognitive effects - introducing science achievement as a control variable

   b) understanding cultural profiles in math teaching

5. Combining TIMSS/8 and PISA, understanding the difference

6. Some suggestions for research stratgies in IEA

Good, Wiley & Florez (2009): Effective Teaching: An Emerging Synthesis.

Opportunity to learn
    → available time, degree of student involvement
Curriculum alignment → visible and coherent plan
Goal-oriented assessment → focus on what is important

Proactive and supportive classrooms → caring communities
Appropriate expectations → help students to exceed

Coherent content → sufficient depth
Thoughtful discourse
Scaffolding students' ideas and task involvement
    → understand at a higher level)
Practice/application (→ concepts in diverse contexts)

Borrowing from psychological research

Process-Product-Research; Behavioral learning theory

→ **Classroom Management, Clarity & Structure**

Humanistic pedagogy and psychology;

Self determination Theory (Deci & Ryan)

→**Supportive climate & Structured learning environment**

Cognitive Theory (e.g. Brown 1997, Mayer 2004);

concepts from (moderate) constructivism

→ **Cognitive Activation & Deep Content**

Pianta & Hamre: Classroom oberservation scales.(CLASS)

- Classroom organization
- Emotional support
- Instructional support

Tschannen-Moran, M. & Woolfolk Hoy, A. (2001):

Ohio teacher efficacy scales (OSTES)

- Efficacy for classroom management
- Efficacy for student engagement
- Efficacy for instructional strategies

DESI Study (German National language Survey) Klieme et al. (2008) Student Questionnaires

Construct	Item Example	Alpha
Structure	At the end of a lesson, the teacher summarizes main issues	.79
Support	My teacher advises me how to improve.	.86
Cognitive Activation	My Teacher stresses that our writing should be gramatically correct	.89

DESI Study
(German National language Survey)
Klieme, Steinert & Hochweber (2010)
Hochweber, Steinert & Klieme (in press)

Standardized effects (controlling for individual and school-level context)
In ninth grade (440 classrooms, 11.000 students)

	German		English	
	Achievement	Interest	Achievement	Interest
Cognitive Activation	.242	.073	.150	.001
Structure	-,147	.188	-.069	.113
Support	.050	.239	.030	.303

For German, effects remain significant if intake is controlled for.

Attention: Quality dimensions are explanatory constructs, NOT "methods" or potential experimental treatments

Rank	Variable	Studies	Effect size
56	**Discovery learning**	205	**.31**
40	**Questioning**	214	**.49**
39	**Cooperative learning**	2285	**.49**
28	**Peer tutoring**	767	**.55**
22	**Worked-out examples**	62	**.57**
18	**Training of learning skills**	656	**.59**
11	**Training of self verbalisation**	92	**.67**
8	**Informative Feedback**	1276	**.72**

Consequences for LSA

1. The notion of OTL should be revisited. What makes education effective is the quality of teaching and learning, not (only) the subject matter per se.

2. The quality of teaching can be validly assessed by students, teachers, and/or observers. There is even evidence that at least „structure/classroom climate" and „support" can be measured cross-culturally (scalar invariance).

3. LSA can be used to **describe** teaching and learning, its **variation** and **distribution** (→ equity ?) based on
(a) subject matter coverage (OTL in the traditional sense)
(b) quality dimensions (structure, support, cognitive activation)
(c) specific practices („treatments").

4. Large Scale assessment data can be used to **model the relationships** with student attainment, cognitive skills, and non-cognitive outcomes.
The explanatory power (for quality dimensions) or causal interpretation (for treatments) depends on the availability of relevant control variables (=approximation for pre-knowledge).

5. International LSA can be used to test the universality vs culture-specificity of teaching quality and effectiveness. (→ Research aim of IEA !!!)

Outline

1. Current challenges of international large scale assessment

2. Teaching in IEA studies: From SIMS to TIMSS Video

3. Theoretical approaches and findings on „teaching quality" ->
   OTL (content coverage), effective methods, process quality

4. Explorations of Teaching practices, quality, and effectiveness
   based on TIMSS 2003, grade 8 Mathematics data →
   a) three-level HLM of cognitive and non-cognitive effects -
      introducing science achievement as a control variable
   b) understanding cultural profiles in math teaching

5. Combining TIMSS/8 and PISA, understanding the difference

6. Some suggestions for research stratgies in IEA

How much do you agree with these statements about your school?

*Fill in **one** circle for each line*

		Agree a lot	Agree a little	Disagree a little	Disagree a lot
a)	I like being in school	①	②	③	④
b)	I think that students in my school try to do their best	①	②	③	④
c)	I think that teachers in my school care about the students	①	②	③	④
d)	I think that teachers in my school want students to do their best	①	②	③	④

Teaching (?) quality:
- support

How often do you do these things in your mathematics lessons?

*Fill in **one** circle for each line*

		Every or almost every lesson	About half the lessons	Some lessons	Never
a)	We practice adding, subtracting, multiplying, and dividing without using a calculator	①	②	③	④
b)	We work on fractions and decimals	①	②	③	④
c)	We interpret data in tables, charts, or graphs	①	②	③	④
d)	We write equations and functions to represent relationships	①	②	③	④
e)	We work together in small groups	①	②	③	④
f)	We relate what we are learning in mathematics to our daily lives	①	②	③	④
g)	We explain our answers	①	②	③	④
h)	We decide on our own procedures for solving complex problems	①	②	③	④
i)	We review our homework	①	②	③	④
j)	We listen to the teacher give a lecture-style presentation	①	②	③	④
k)	We work problems on our own	①	②	③	④
l)	We begin our homework in class	①	②	③	④
m)	We have a quiz or test	①	②	③	④
n)	We use calculators	①	②	③	④

How often do you do these things in your mathematics lessons?

*Fill in **one** circle for each line*

	Every or almost every lesson	About half the lessons	Some lessons	Never
a) We practice adding, subtracting, multiplying, and dividing without using a calculator	①	②	③	④
b) We work on fractions and decimals	①	②	③	④
c) We interpret data in tables, charts, or graphs	①	②	③	④
d) We write equations and functions to represent relationships	①	②	③	④
e) We work together in small groups	①	②	③	④
f) We relate what we are learning in mathematics to our daily lives	①	②	③	④
g) We explain our answers	①	②	③	④
h) We decide on our own procedures for solving complex problems	①	②	③	④
i) We review our homework	①	②	③	④
j) We listen to the teacher give a lecture-style presentation	①	②	③	④
k) We work problems on our own	①	②	③	④
l) We begin our homework in class	①	②	③	④
m) We have a quiz or test	①	②	③	④
n) We use calculators	①	②	③	④

DIPF
Educational Research
and Educational Information

Teaching quality:
- cognitive activation (g, h)

„Treatments":
- group work (e)
- lecturing (j)

Math content dimensions:
- pure math (a,b,d)
- applications (c, f)

Predictor	Dependent: Valueing Math
Student:	
Classroom:	
Cogn. Activation	.20 ***
Moderation: Country level Cogn. Activation	.07
Support	.21
Moderation: Country level Support	.23 **

Predictor	Dependent: Valueing Math	Dependent: Math Achievement
Student:		
Classroom:		
Cogn. Activation	.20 ***	42.93 ***
Moderation: Country level Cogn. Activation	.07	1.25
Support	.21	11.51
Moderation: Country level Support	.23 **	-43.64

Predictor	Dependent: Valueing Math	Dependent: Math Achievement	Dependent: Math Achievement
Student: Science Ach			.498 ***
Classroom: Science Ach			.964 ***
Cogn. Activation	.20 ***	42.93 ***	11.37 ***
Moderation: Country level Cogn. Actvation	.07	1.25	7.36
Support	.21	11.51	-3.14
Moderation: Country level Support	.23 **	-43.64	-14.95

Predictor	Lecturing
Student:	
N of books	9.42***
Classroom:	
N of Books	54.88 ***
Science Ach	
Lecturing	20.33 **
Group work	
Moderation: Country level Math Ach	0.03

Predictor	Lecturing	
Student:		
N of books	9.42***	4.01 ***
Science Ach		0.48 ***
Classroom:		
N of Books	54.88 ***	8.60 ***
Science Ach		0.91 ***
Lecturing	20.33 **	4.06 ***
Group work		
Moderation: Country level Math Ach	0.03	0.05 ***

Predictor	Lecturing		Group work
Student:			
N of books	9.42***	4.01 ***	9.42 ***
Science Ach		0.48 ***	
Classroom:			
N of Books	54.88 ***	8.60 ***	54.22 ***
Science Ach		0.91 ***	
Lecturing	20.33 **	4.06 ***	
Group work			-14.53 ***
Moderation: Country level Math Ach	0.03	0.05 ***	.05 *

Predictor	Lecturing		Group work	
Student:				
N of books	9.42***	4.01 ***	9.42 ***	4.01 ***
Science Ach		0.48 ***		0.48 ***
Classroom:				
N of Books	54.88 ***	8.60 ***	54.22 ***	8.21 ***
Science Ach		0.91 ***		0.92 ***
Lecturing	20.33 **	4.06 ***		
Group work			-14.53 ***	-1.44
Moderation: Country level Math Ach	0.03	0.05 ***	.05 *	-.12 ***

Summary of findings (TIMSS 03)

- „**Cognitive activation**" had a positive effect both on cognitive and affective outcomes in mathematics across countries. (Although there was considerable variation in slopes between countries.)

- „**Support**" had a positive effect on students' valuing mathematics – but only in countries that generally had a supportive culture.

- If **science achievement** (as a proxy for general academic ability) is controlled for, the positive effect of „cognitive activation" on mathematics achievement still exists.

- **Lecturing** was linked to better mathematics achievement, and more so in high achieving countries – although the effect is rather tiny. **Group work** seemed to be negatively linked with mathematics achievement. However, by using a control for academic ability, we see that this effect only holds in high achieving countries.

- Generally, there is substantial variation of slopes between countries, and effect sizes are small. This is just an explorational starting point !
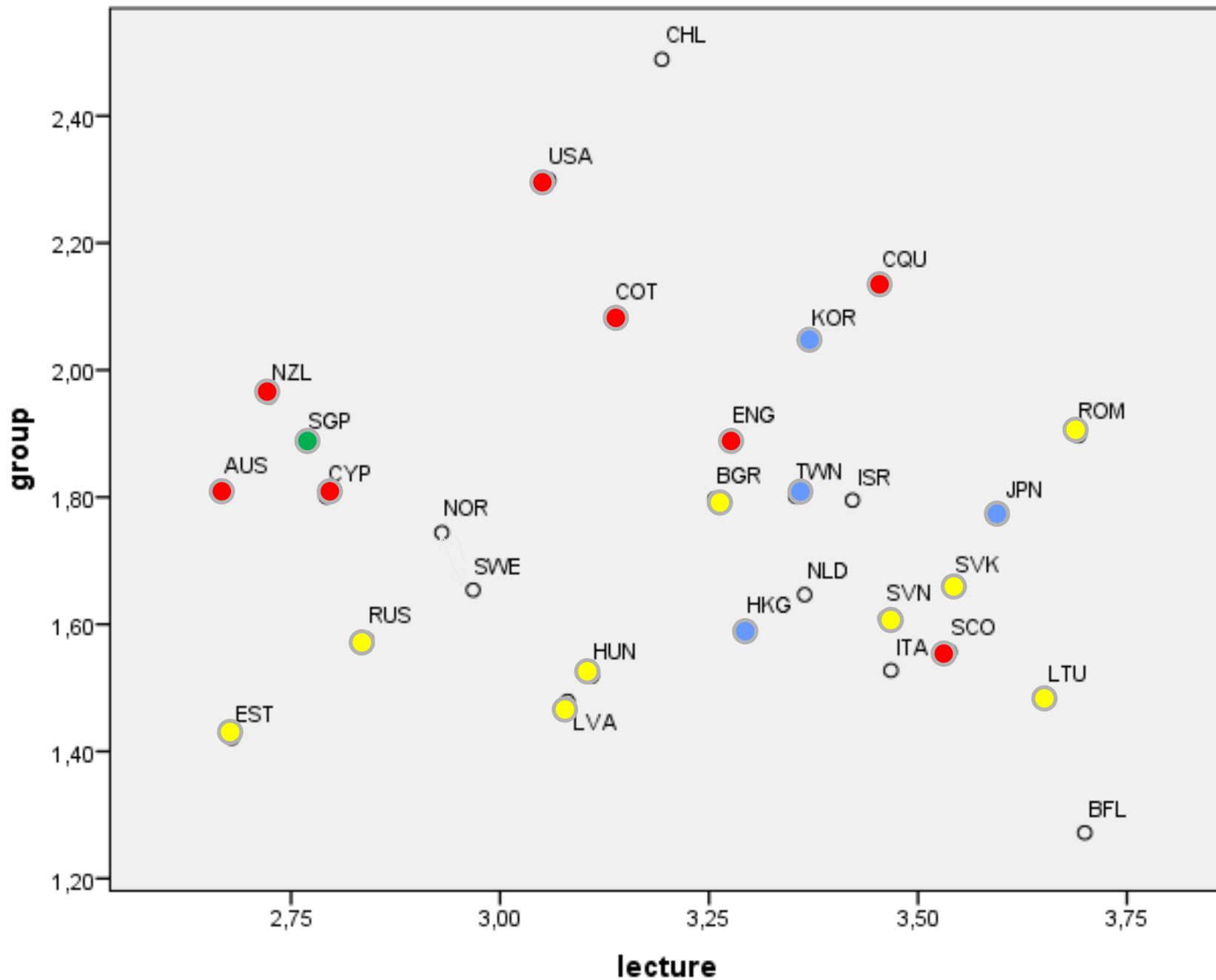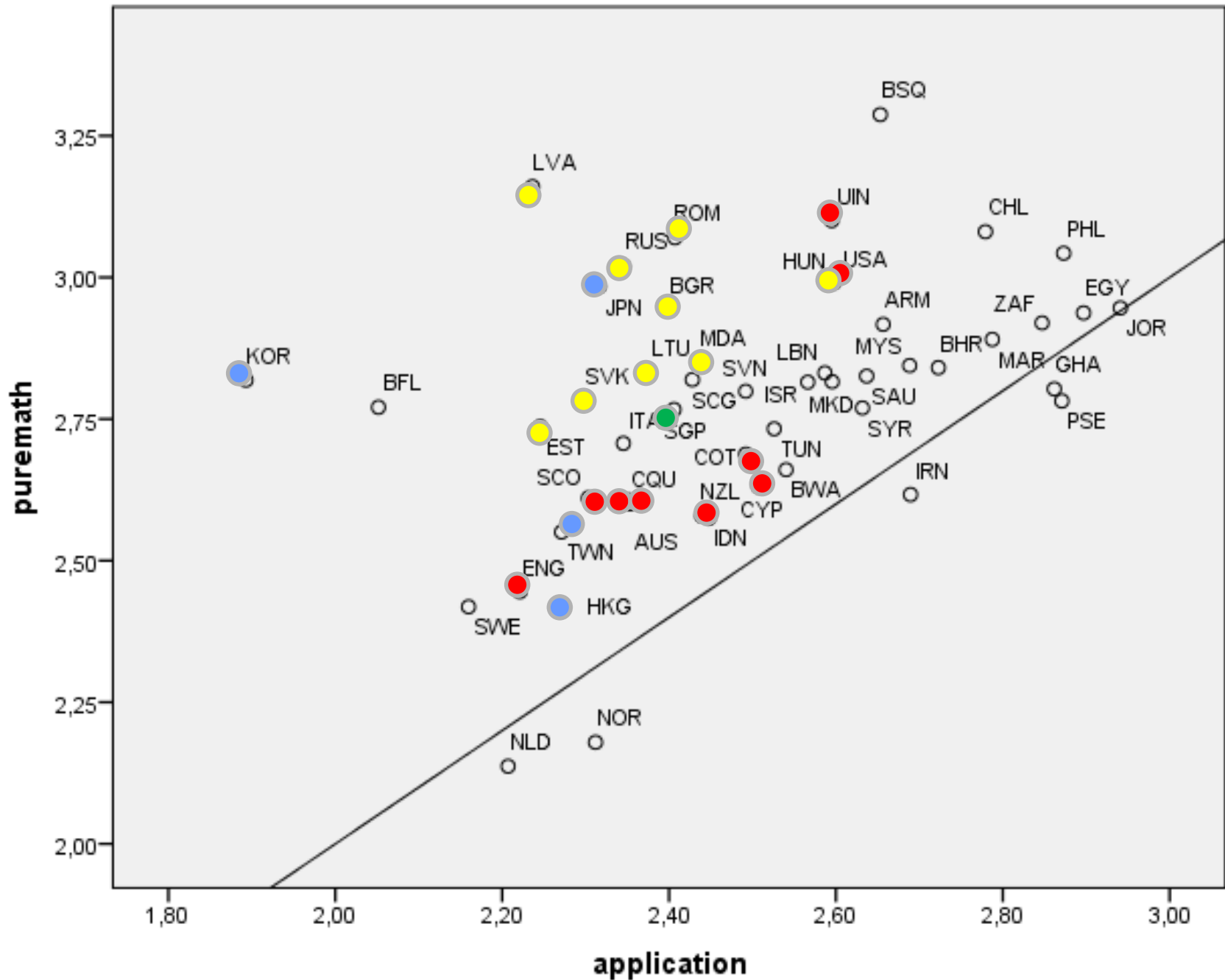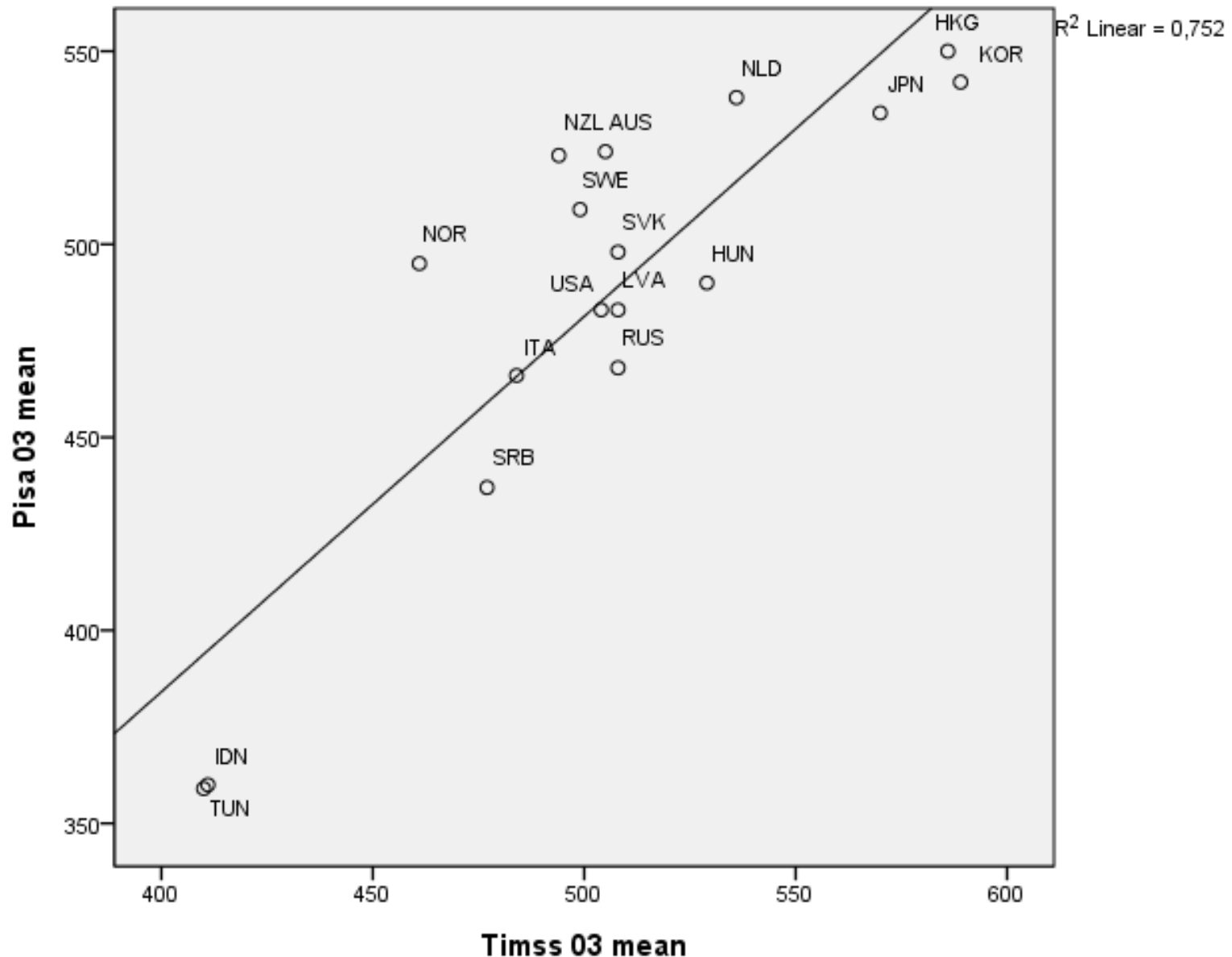
Outline

1. Current challenges of international large scale assessment

2. Teaching in IEA studies: From SIMS to TIMSS Video

3. Theoretical approaches and findings on „teaching quality" -> OTL (content coverage), effective methods, process quality

4. Explorations of Teaching practices, quality, and effectiveness based on TIMSS 2003, grade 8 Mathematics data →

   a) three-level HLM of cognitive and non-cognitive effects - introducing science achievement as a control variable

   b) understanding cultural profiles in math teaching

5. Combining TIMSS/8 and PISA, understanding the difference

6. Some suggestions for research stratgies in IEA

Explaining the gap between PISA and TIMSS (Klieme et al., 2013)

Regression on Country level (n=17): R² = .93

Dependent: PISA 2003

Predictor	
TIMSS 2003	. 86
Pure Math experience (as measured in PISA 2012 Field Trial)	- 87
TIMSS mean age	- 30

Outline

1. Current challenges of international large scale assessment

2. Teaching in IEA studies: From SIMS to TIMSS Video

3. Theoretical approaches and findings on „teaching quality" -> OTL (content coverage), effective methods, process quality

4. Explorations of Teaching practices, quality, and effectiveness based on TIMSS 2003, grade 8 Mathematics data →

    a) three-level HLM of cognitive and non-cognitive effects - introducing science achievement as a control variable

    b) understanding cultural profiles in math teaching

5. Combining TIMSS/8 and PISA, understanding the difference

6. Some suggestions for research stratgies in IEA

Some suggestions for research stratgies in IEA (1)

The study design and the context assessment mostly determine the explanatory power of LSA.

Stakeholders should understand that context assessment needs more effort (and response time !).

Strong theoretical frameworks are available.

Refresh the concept of OTL to include quality factors: „Implemented curriculum" is more than content matter: it is the „enacted instructional regime" (Raudenbush).

Develop innovative measures: rating scales based on generalized statements and frequency of practices; observational methods, artifacts.

Some suggestions for research stratgies in IEA (2)

The IEA design principles (classroom based, multi-subject, curriculum-related) offer a good foundation. Try add-ons (at least for subgroups of countries): additional context scales, improved control variables (incl. a general reasoning test), multiple classrooms within schools, longitudinal add-ons (short term for teaching effects, medium-term for achievement growth, long-term for school cahnge, very long term for predictive validation). Adjust test design accordingly.

Develop theory and modeling techniques for the system level (trend).

Let research drive IEA studies, not just policy making !