



**DIPF**

Educational Research  
and Educational Information



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

# Coordination of assessment at the national and state levels in Germany

Eckhard Klieme<sup>1</sup> & Knut Schwippert<sup>2</sup>

<sup>1</sup>German Institute for International Educational Research (DIPF)  
Frankfurt am Main, Germany

<sup>2</sup>University of Hamburg, Germany

IEA General Assembly

Lisbon, Portugal, October 7-10, 2013



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions



**DIPF**

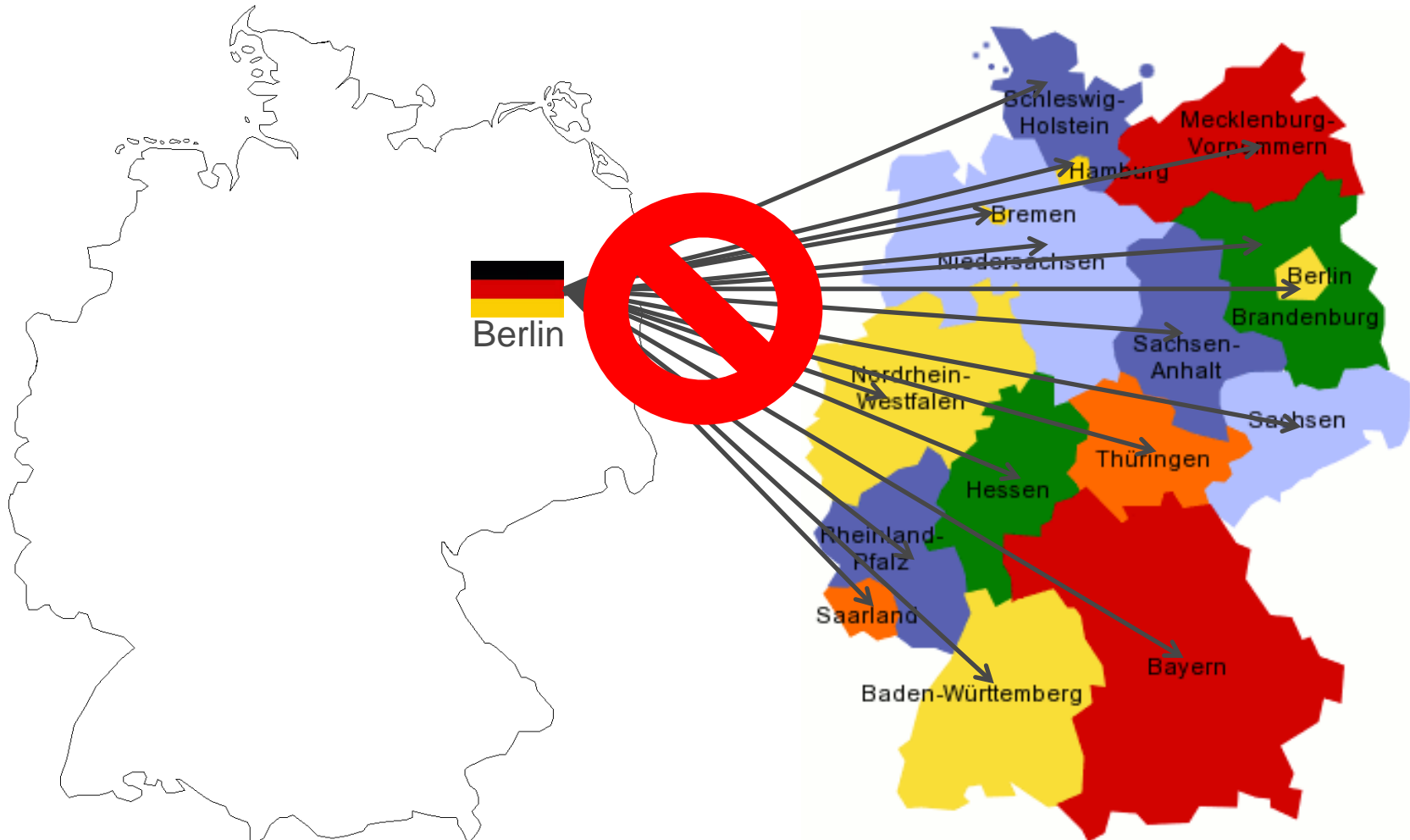
Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Germany (Bund) and it's 16 Federal States (Länder)



<http://juergen-schwarz-online.de/html/reisewetter.html>

<http://www.loreleyinfo.de/deutschland-germany.php>  
Bundesländer in Deutschland © David Liuzzo



## Länder + Bund

Steering Committee for National Monitoring of the Educational System (with scientific advisory board)

→ runs:

**ZIB** (Centre for international large-scale studies): founded in 2009 to implement international studies, especially PISA, and foster research. Supports research groups at TUM (Munich), DIPF (Frankfurt), IPN (Kiel), and several universities

## Länder only

Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK)

→ runs:

**IQB** (National Institute for School Quality Monitoring) at Humboldt University of Berlin

to develop national standards and coordinate national assessments

**IEA-Studies** run by „IEA Deutschland e.V.“, a network of researchers and institutes with national study centers at TU Dortmund (PIRLS, TIMSS, ICILS), Univ. of Essen (ICCS), Humboldt Univ. Berlin (TEDS-M)

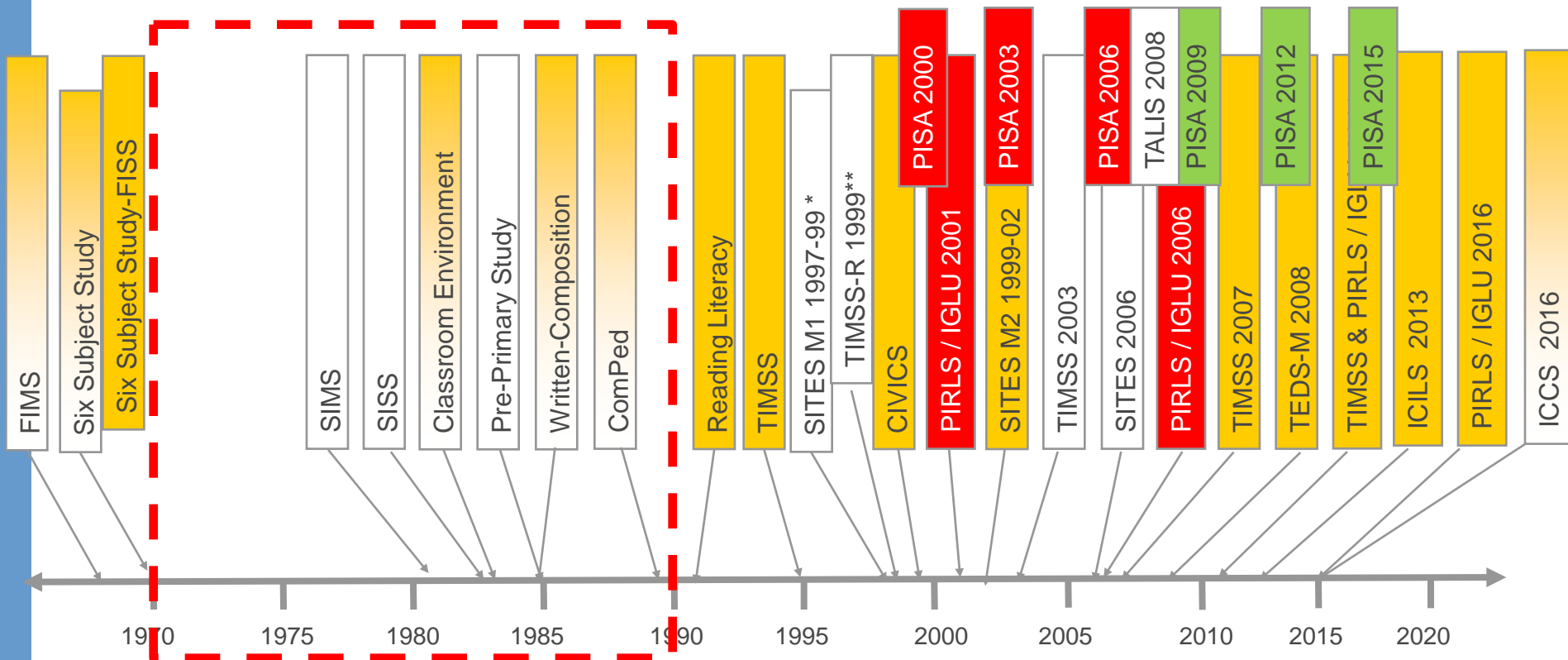
## Academic Network for Educational Assessment

also including, e.g., a Priority Research Program on Educational Measurement



# International Assessments: German Participation

with enhancements for intra-national comparison



# International Assessments: German Assessment Design

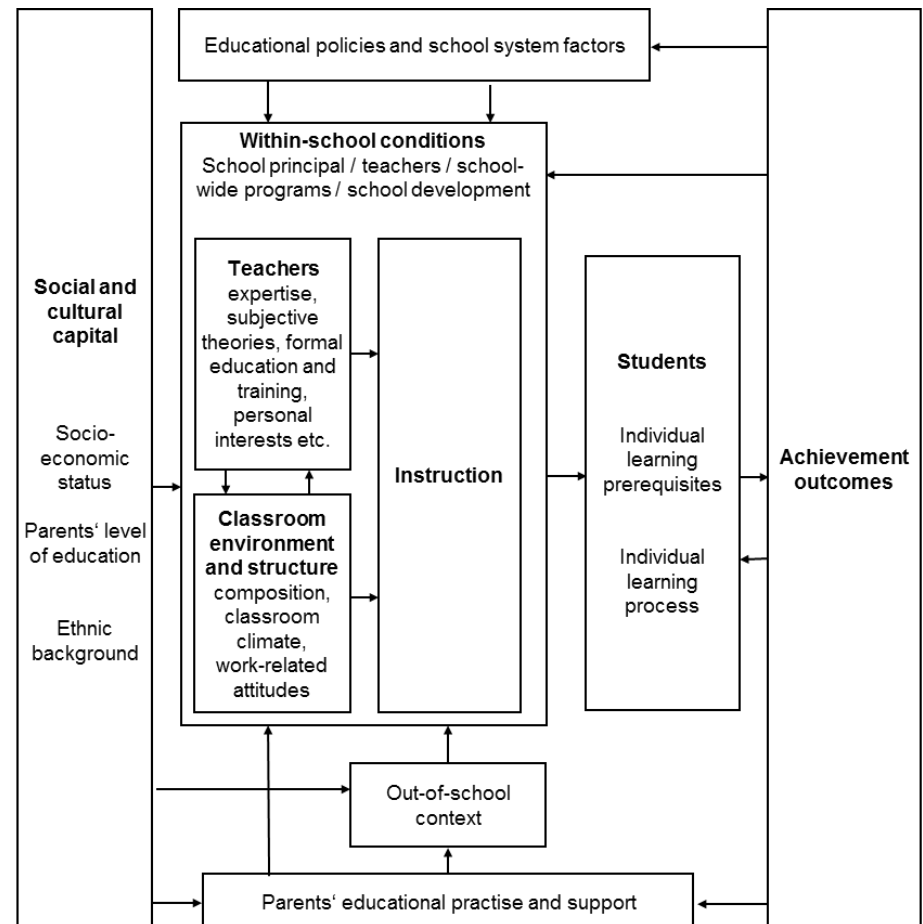
with enhancements / extensions to address specific research questions

## PIRLS Germany Framework

Modified production model of  
H.J. Walberg

### Theories on different levels

- Sociology  
(Bourdieu, Bourdon)
- Psychology  
(Krapp, Deci & Ryan)
- School  
(Fend)
- Teaching  
(Helmke, Weinert)



# National Standards-based Assessment („Ländervergleich“): Schedule and combination with international studies

**Overview: Assessments and preportin between 2006 and 2017;**

**Every 5 years in primary school, every 6 years in seconday school I**

Year	Assesment				Reporting			
	PIRLS	TIMSS	PISA	Ländervergleich	PIRLS	TIMSS	PISA	Ländervergleich
2006	✓		✓					
2007		✓			○		○	
2008					■	○	■	
2009			✓	✓ (G; E; F)				
2010							○	● (G; E; F)
2011	✓	✓		✓ (PS)				
2012			✓	✓ (M; Sc)	○	○		● (PS)
2013							○	● (M; Sc)
2014								
2015		✓	✓	✓ (G; E; F)				
2016	✓			✓ (PS)		○	○	● (G; E; F)
2017					○			● (PS)

**Note: International Comparison; Ländervergleich (PISA or PIRLS); Ländervergleich on basis of national standards;  
PS: Primary School Standards Mathematics and German, G: German Sec I; E: English Sek I; F: France Sek I;**



## **National Assessment for School-level Evaluation (VERA)**

Every year, all students in

3<sup>rd</sup> grade (Primary Schools) – German or Mathematics

8<sup>th</sup> grade (Lower Secondary Schools) German, Math, or EFL/FFL

Local Administration and Scoring, Central Analysis and Reporting

Feedback to schools and teachers,

Sometimes individual feedback to students and parents,

sometimes results published on school level

Mostly low stakes for students

No reporting on state (Länder) level,

less feasible for research



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings **(Slides prepared by Olaf Köller, IPN)**
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions

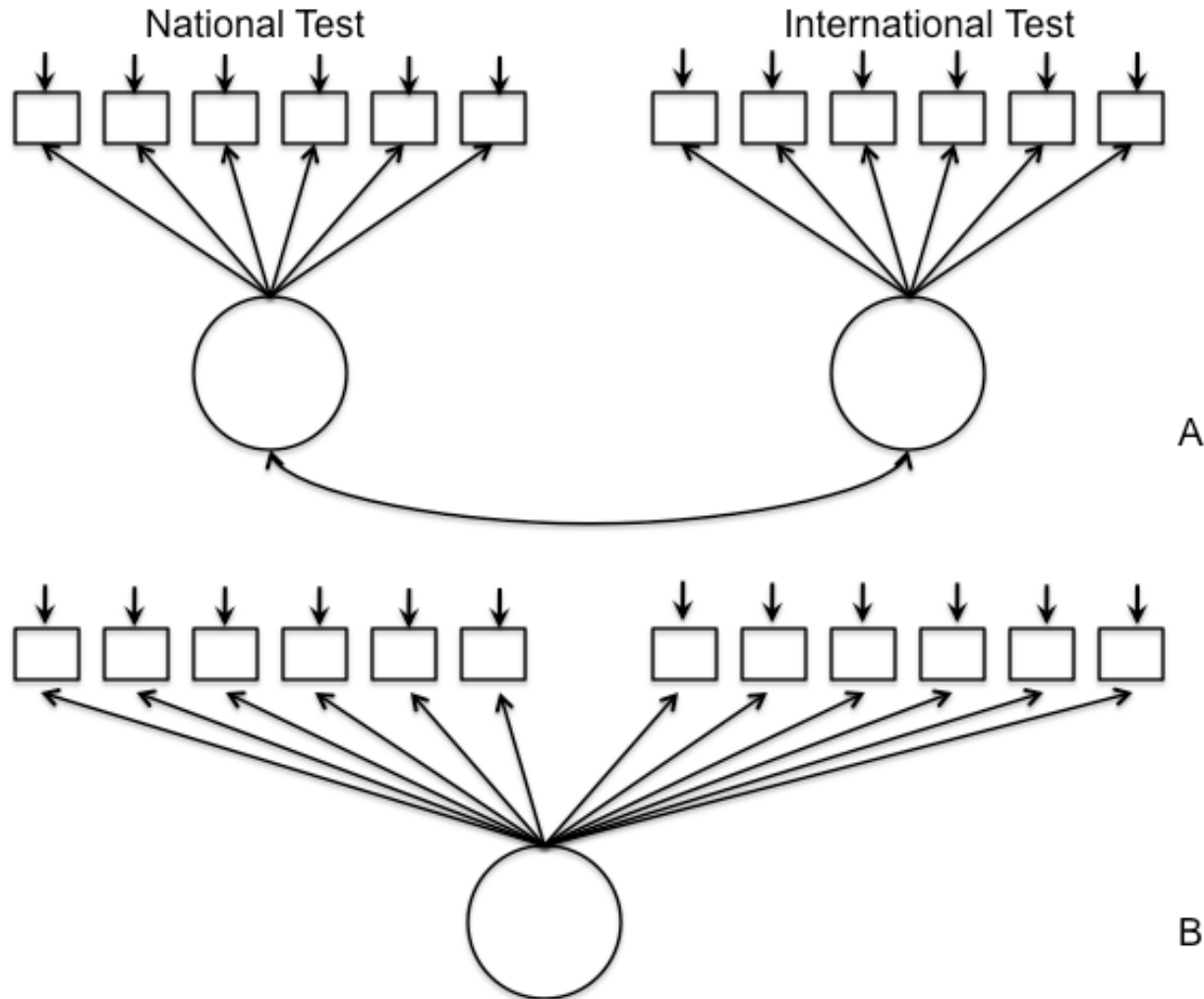


# Research Questions

Strong political as well as scientific pressure to link studies (see e.g., similar activities in USA, where NAEP 8 and TIMSS have been linked in 2011)

- I. Different tests, different constructs?
- II. Different tests, different proficiency level models?
- III. Can we use national tests to assess our students on international scales and vice versa?

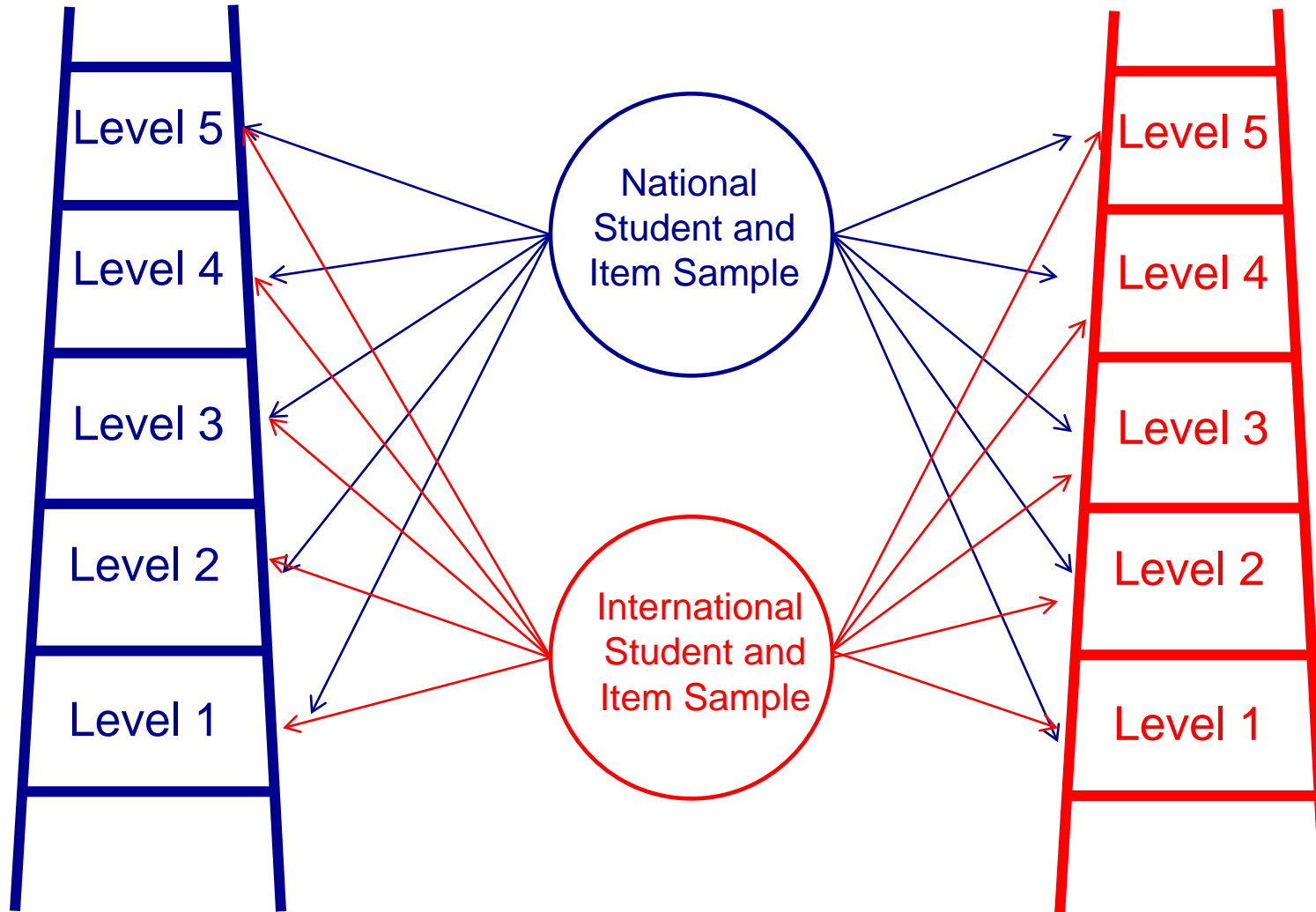
# Research Question I: Graphical Illustration



# Research Questions II and III: Graphical Illustration

National Proficiency Levels

International Proficiency Levels





# Linking PIRLS 2006 and German National Assessment\*

- Linking the international PIRLS reading scale with our national reading scale (large field trial)
- Sample:  $N$  4.728 4<sup>th</sup> graders (50.1% female)
- First day: PIRLS (multi-matrix-design; 80 minutes)
- Second day: National items on reading, listening, spelling and writing (multi-matrix-design; 80 minutes)



# Findings

- Correlation of  $r = .84$  (estimate from two-dimensional analysis)
- Population parameters
  - PIRLS original:  $M = 548$ ,  $Sd = 68$ 
    - National Items:  $M = 547$ ,  $Sd = 55$
- Due to differences in standard deviations different proportions of students assigned to proficiency levels





## Findings

- Correlation of  $r = .84$  (estimate from two-dimensional analysis)
- Population parameters
  - PIRLS original:  $M = 548$ ,  $Sd = 68$
  - National Items:  $M = 547$ ,  $Sd = 55$
- Due to differences in standard deviations different proportions of students assigned to proficiency levels

## Conclusions

- RQ I: National test strongly related to, but not identical to PIRLS
- RQ II: Different proficiency models (levels)
- RQ III: Estimation of mean average on PIRLS scale through national assessment would be possible (e.g., for Länder)



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

# Linking TIMSS/PIRLS 2011 and National Assessment

National Assessment  
German/Math

4<sup>th</sup> graders (1 class per  
school)

1.300 schools

TIMSS/PIRLS

4<sup>th</sup> graders (1 class per  
school)

201 schools

$r = .92$  for Mathematics (Pietsch et al., 2013)



## Linking TIMSS/PIRLS 2011 and National Assessment

National Assessment  
German/Math

4<sup>th</sup> graders (1 class per  
school)

1.300 schools

TIMSS/PIRLS

4<sup>th</sup> graders (1 class per  
school)

201 schools

$r = .92$  for Mathematics (Pietsch et al., 2013)

## Linking PISA with National Assessment

$r = .85$  for Reading (Jude et al., 2013)

$r = .95$  for Mathematics (Hartig & Frey, 2012)



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions



DIPF

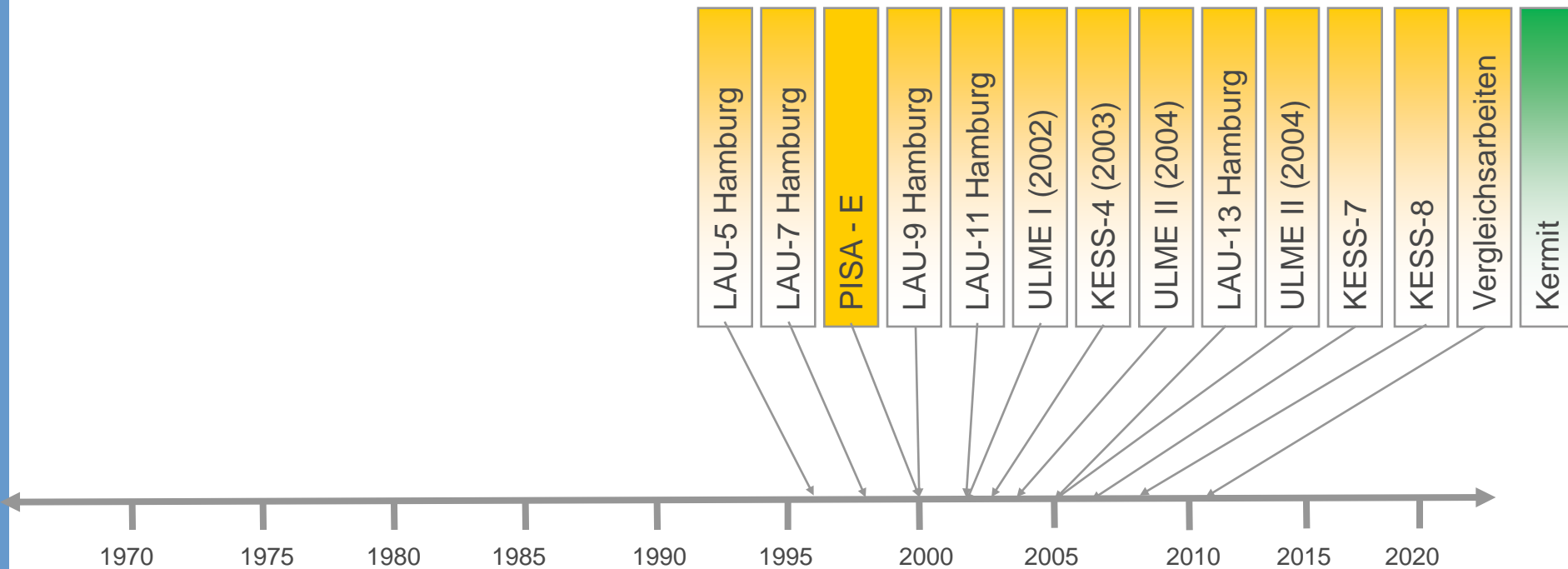
Educational Research  
and Educational Information



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Regional System Monitoring Studies in Hamburg





DIPF

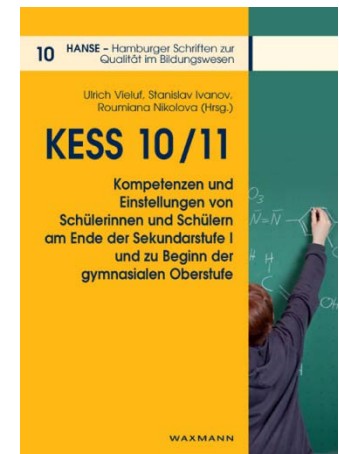
Educational Research  
and Educational Information



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

# Some need for information .... (in Hamburg)



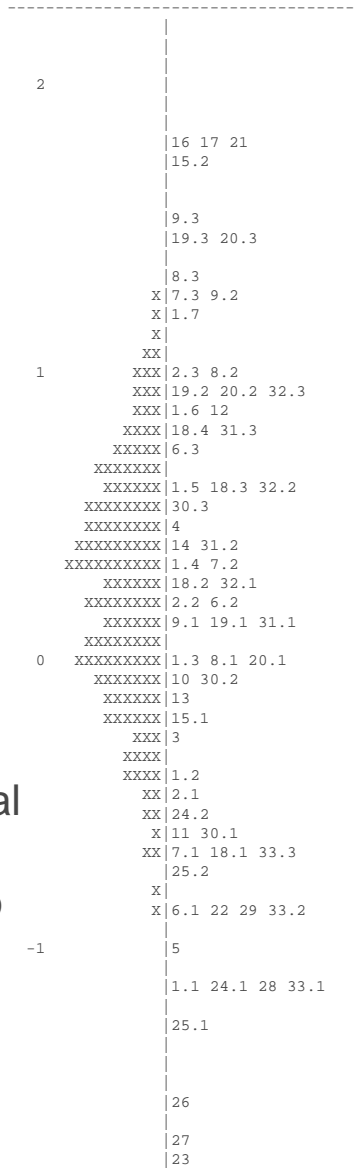
# Social Index (Hamburg)

Construct based on:

- social capital
- economic capital
- cultural capital
- migration background

Individual answers are aggregated on school level.

Most variables are well established in international large scale surveys – but the list of items needed to be extended to address specific (regional) characteristics



- Household income
- ISEI
- Loan mower
- Second car
- Access to the internet
- Books at home
- Children books
- Father's highest education
- Mother's highest education
- Newspaper
- Learning Software
- Fairy Tales
- Items of art
- Visiting Theater
- Visiting Museum or Art-Exhibition
- Visiting Opera, Ballet or classical concerts
- Visiting Rock-, Pop- or Jazz-concerts
- Visiting Sport events
- Visiting Cinema
- Visiting "Stadtteilfest, Volksfest oder Jahrmarkt"
- Visiting Presentation or Readings (books)
- Migration status of parents
- My parents do always know where I am after school
- I talk with my parents about my plans with I have with my friends
- When I am leaving home, my parents ask where I am going.
- When I meet my friends, I tell my parents with whom I will meet.
- I know my child's friends
- I always know, with which friends my child is out.
- I talk with my child about the plans it has with it's friends
- My parents pay a lot of attention about how much time I spend for homework
- My parents like, if I make my homework always at the same time.
- My parents insist, that I spend a specific time for reading
- My parents want that do first my homework before I meet my friends

## Social Index (Hamburg)

The index is used for allocation of resources and decision making on educational programs, e.g.

- reduction of class size in disadvantaged schools
- more time for language teaching depending on class size and social index
- additional early language assessment at 4½
- provision of all day care and education
- Inclusion – benefits depending from social index
- more administrative support in disadvantaged schools
- support for private schools depending on social index





**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## **Determination of competencies in Hamburg (Kompetenz Ermittlung) since 2013**

**KERMIT** takes place in grades:

2: German, mathematics

3: German, mathematics (= Lernstand 3)

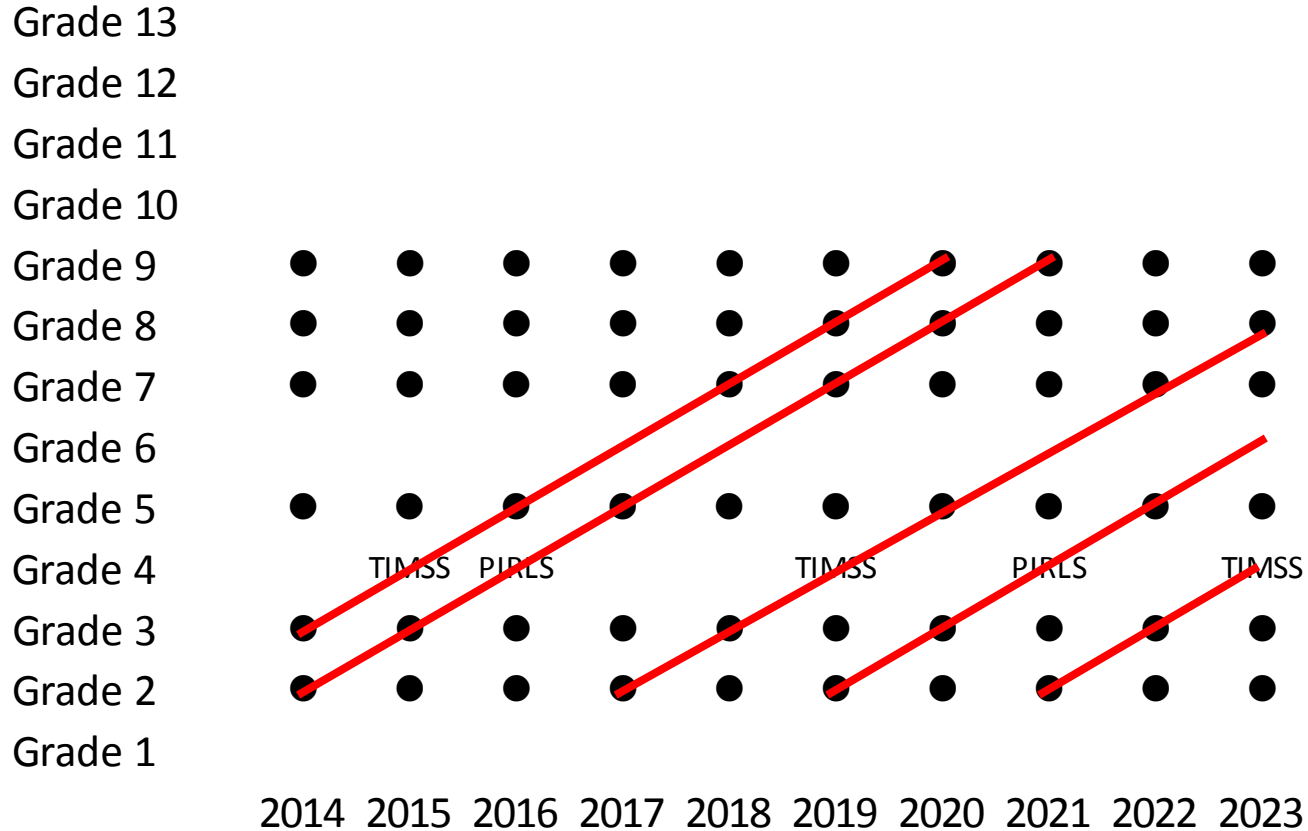
5: German, mathematics, science, English

7: German, mathematics, science, (since 2013 also English)

8: German, mathematics, English / French (= Lernstand 8)

9: German, mathematics, science, English

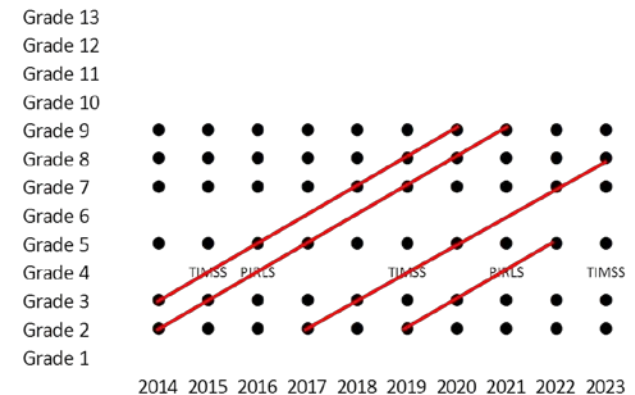
## Determination of competencies in Hamburg (Kompetenz Ermittlung) since 2013



## Determination of competencies in Hamburg (**K**ompetenz **E**rmittlung) since 2013

### Consequences

- good longitudinal data
- in-depth monitoring
- 2 assessments in primary school every year
- 4 (!) assessments in secondary schools every year





**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## Overview

1. Monitoring and Evaluation in German schools:  
Decision making, Administrative Structure, Studies
2. Linking National and international Assessments:  
Empirical findings
3. Assessment strategies at the state level –  
The example of Hamburg
4. Conclusions



# Why Combining International and National Studies

## International

### Advantages

- International Comparison
- State of the art achievement-assessments

### Disadvantages

- Reduced common denominator
- Limited scope of background questionnaires

## National

### Advantages

- In-depth view
- Focus on specific issues within states

### Disadvantages

- Limited Comparisons
- Limited access to data for independent researchers



# Why Combining International and National Studies

## Facing Problems

- displacement of independent research by overburden of schools
- feedback strategies to
  - politicians
  - administration
  - school (principal / teacher / parents)
- narrowed (national) view on educational system

## Perspectives

- Overarching coordination of (nat.) assessments (sampling)
- Addressing the demand of information from
  - politicians
  - administration
  - school (principal / teacher / parents)
  - fundamental research
- broader (national and international) view on educational system



**DIPF**

Educational Research  
and Educational Information



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

## **Mission statement** ([www.iea.nl](http://www.iea.nl))

The IEA Secretariat and Data Processing and Research Center, together with the association's membership, carry out comparative research studies in education. Through its comparative research and assessment projects, IEA aims to:

- provide international benchmarks to assist policy-makers in identifying the relative strengths and weaknesses of their education systems
- provide high-quality data to increase policy-makers' understanding of key school- and non-school-based factors that influence teaching and learning
- provide high-quality data that will serve as a resource for identifying areas of concern and action, and for preparing and evaluating educational reforms
- develop and improve the capacity of education systems to engage in national strategies for educational monitoring and improvement
- contribute to the development of a worldwide community of researchers in educational evaluation.

# Perspectives for IEA Studies

Continue to strengthen achievement assessments

Enhance Background Questionnaire:

- Theoretical approach
- Empirical approach

Distinguish between

- obligatory (for all countries)
- optional (for interested countries)
- national adaptations (by countries needs)

Longitudinal studies

Process data (teaching)



“We choose to .... and do the other things, not because they are easy, but because they are hard, ... “  
(JFK Sep. 12th 1962)





**DIPF**

Educational Research  
and Educational Information



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

# Thank you for your attention

Eckhard Klieme

Knut Schwippert