

Reporting on Education in the Context of Evidence Based Governance

Prof. Dr. Knut Schwippert
University of Hamburg

4.3.2010

- A brief historical view
- The relation between policy-maker and researcher
- Some key-issues by thinking about “Evidence Based Governance”
- Future perspectives

Principles of scientific management: Key attributes for the management of any organization or activity:

1. Intuitive methods of decision making should be replaced by a more scientific method based on observation and analysis;
 2. Managers should be selected for their positions scientifically be appropriately trained;
 3. Managers should ensure that work is carried out according to established standards and procedures; and
 4. Duties should be allocated so that managers assume responsibility for the key tasks of planning, preparation, and supervision.
- Taylor 1911
- (Ross & Mählick 1990: 68)

“The causes behind the failings of the present educational system and the consequences of new structures and contents can be identified and analysed only by scientific research. Thereby educational research becomes an essential prerequisite for educational reform.”

- Comment to PISA 2000, 2003, 2006 results?
 - Statement of the KMK 2006?
 - Reaction to TIMSS 1995 results?
 - Reaction to IRLS 1991 results?
 - Bericht der Bundesregierung zur Bildungspolitik 1970! (Bildungsbericht '70)
- Husén (1984: 23)



6

Evidence Based Decision

Prof. Dr. Knut Schwippert

- After the TIMSS- (end of 1990'ies) and the PISA- (2000) shock significant efforts have been undertaken to change the “steering” of the educational system
- *Governance* as a special case of “steering” / “advising” of social systems
- From input- over process- to output-oriented governance

7

How do policy-maker and scientists communicate

The model of “Brokerage Agencies”

Prof. Dr. Knut Schwippert

8

Relationship of Researcher and Policy-Makers

Prof. Dr. Knut Schwippert

Groups of Interest:

1. Researchers
2. Policy-Makers/ Planers
3. “Middle man” (professionals who are part of the Ministry and who serve in a liaison function)

The last group rather quickly tends to become socialized into the orientations and opinions prevailing among administrations and/or policy-makers in the Ministry. Therefore it would be justified to distinguish between two main categories only: researcher and policy-makers.

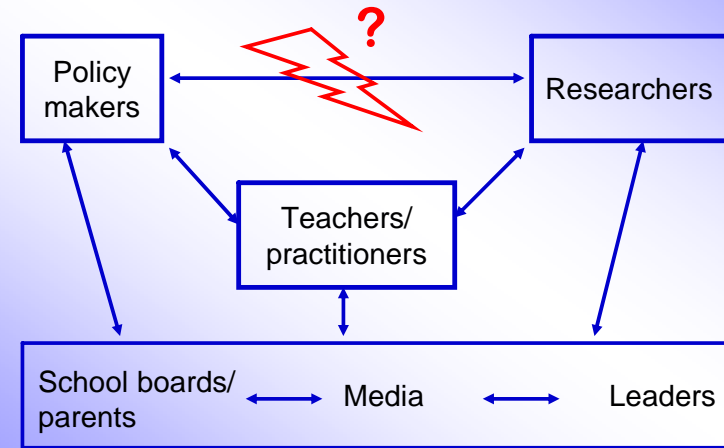
Husén (1984: 8)

Relationship of Researcher and Policy-Makers

“Policy-makers want to use research in order to legitimise a “pre-fabricated” position, only to find that the findings are not supportive or are even used as ammunition by adversaries. This is something that social research, given its fields of inquiry, has to live with and can take with a certain magnanimity, as long as its results are available to all interested parties.”

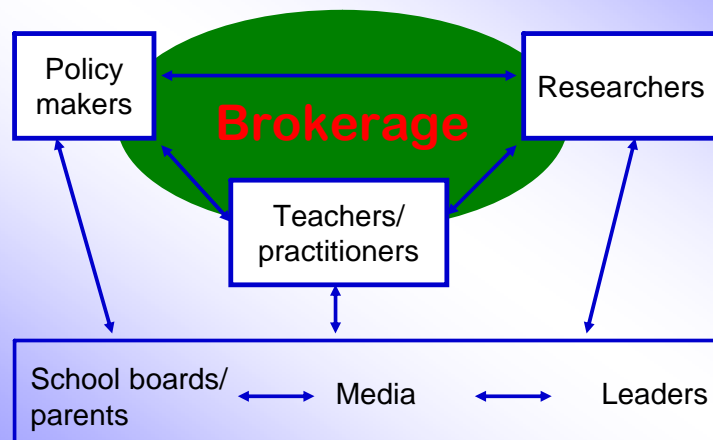
Husén (1984: 14)

Lines of Communication

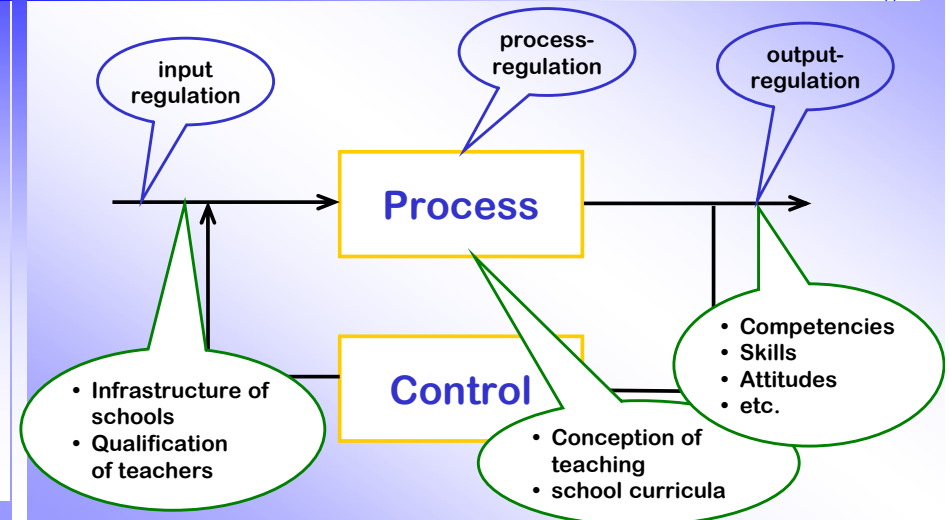


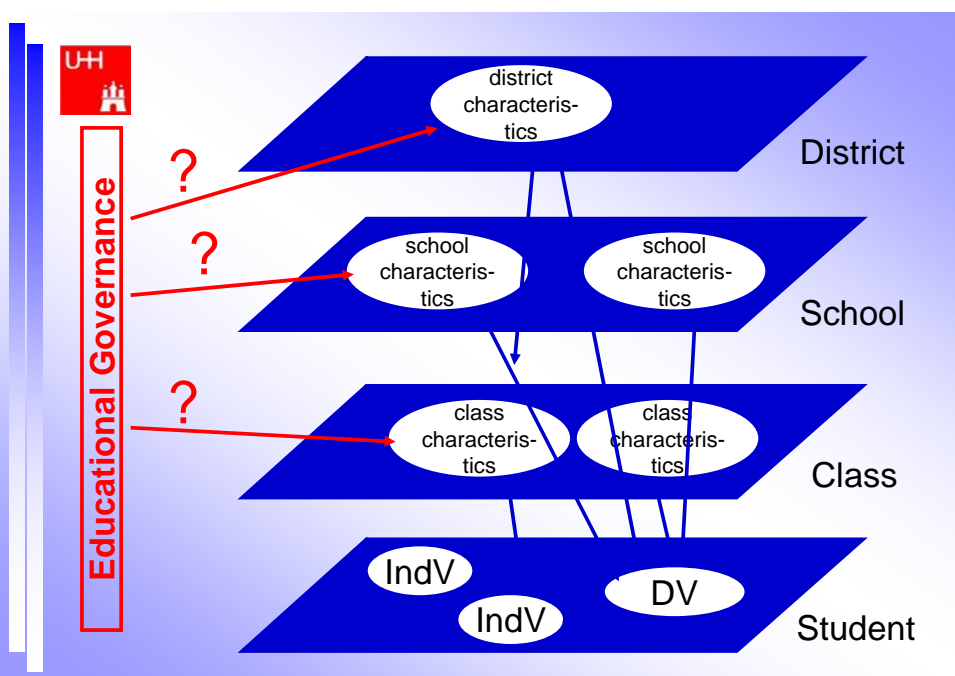
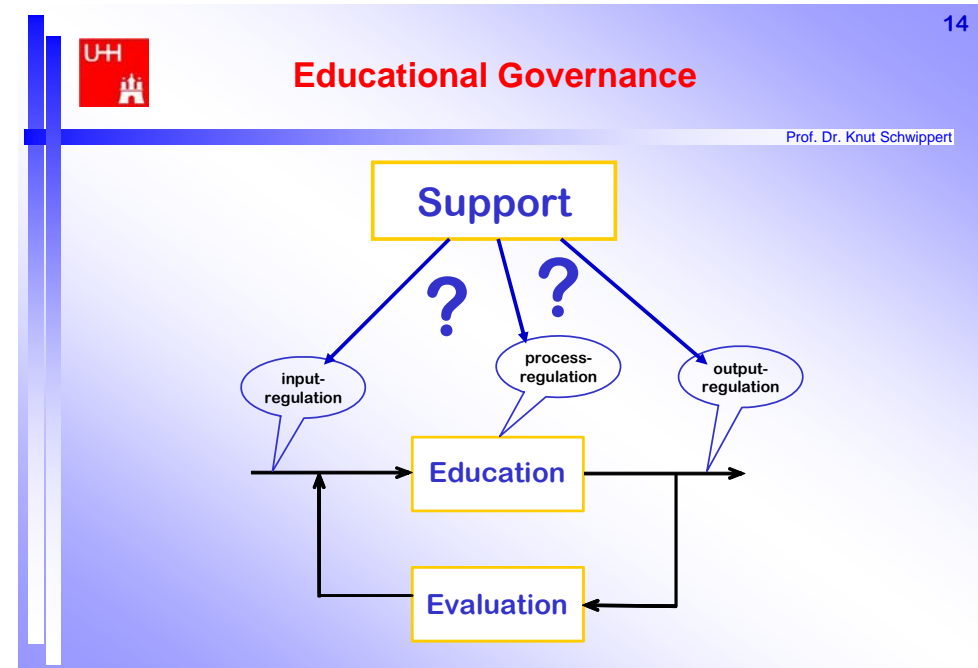
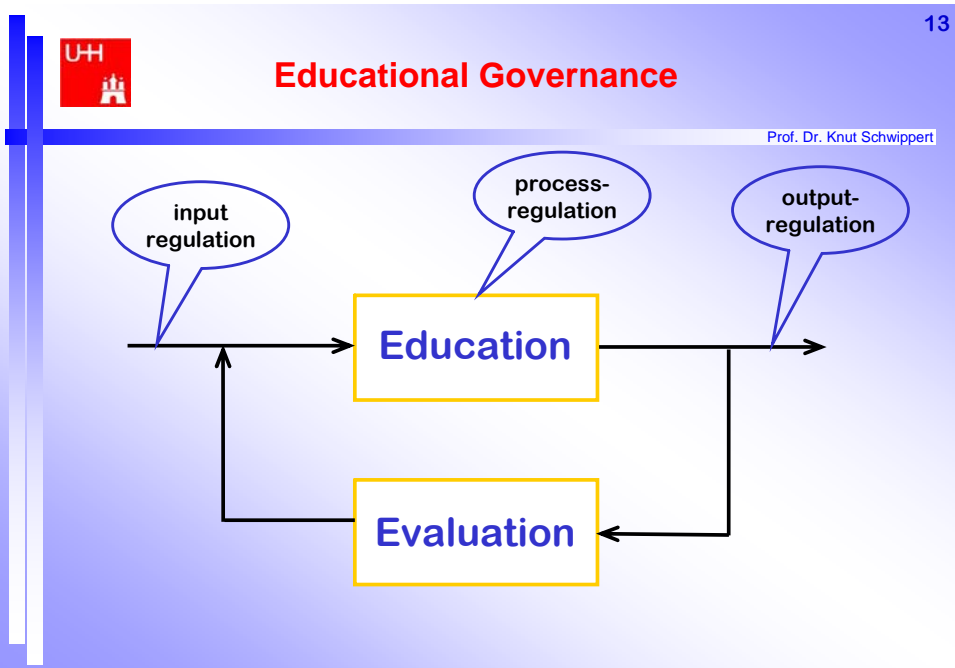
(CER/OECD 2007)

Lines of Communication



Educational Governance





- 16
- UH
- ## Today's Situation in Germany
- Prof. Dr. Knut Schwippert
- At the moment some more attention needs to be paid to what kind of studies (evaluation) gives what kind of evidence!
 - We have to distinguish:
 - International system monitoring studies
 - National (regional) system monitoring studies
 - Longitudinal studies
 - Experimental design studies
 - but why?



Study Design

Klieme, Baumert & Schwippert 2000

Prof. Dr. Knut Schwippert

Design		Targets
(Controlled) Intervention Studies	Research Praxis	Explanation of conditions of effectiveness of educational aims
Longitudinal Studies		Analysis of trajectories of development under consideration of institutional and psychosocial factors
System Monitoring		Description of (a) framework, (b) school and teaching, (c) competencies / attitudes of students, and (d) it's correlations
Evaluation Studies of Schools and School-Programs		Target oriented, activity based examination und benchmark of strength and weakness
Monitoring of Educational Reforms		Examination of initial conditions, consultation and process evaluation



Key-Questions

Prof. Dr. Knut Schwippert

- What is “Evidence” in Educational Reporting?

What works?
or with other words
What travels?

- What will be observed in which context?
- What will be observed and how long?



Some general thoughts about the relation of science and administration

Prof. Dr. Knut Schwippert

Evidence ?

Everybody can use PISA for himself or his/her personal interest. Research in this form can be explained as “neutral of use“ ...

(Tenorth 2004: 62)



The focus of interest

Prof. Dr. Knut Schwippert



Levels of Reporting

Prof. Dr. Knut Schwippert

- **Special situation in Germany:**
The relation between “Bund und Länder”
- **Educational Reporting**
 - on international level (IEA, IIEP, OECD etc pp.)
 - on national level (since 2003 / 2006 + 2008)
 - on regional (federal state) level (since 2006 in seven federal states)
 - on community (city) level
 - on school level



Today's Situation in Germany

Prof. Dr. Knut Schwippert

- Along with the implementation of educational standards, central comparative tests and high stakes tests (certificates) the construction of a system-based external evaluation of schools became a central instrument in the governance of the school system. Although we find in Germany different wordings the core idea is the same in all states: The systematic evaluation for the inventory about the terms, the process and the results of schools lies in the focus.

(Döbert et al. 2008: 63)



Indicators

Prof. Dr. Knut Schwippert

- Indicators in a psychometric context are single information which can be defined and measured directly (indicator) or indirectly (scale).
- Indicators in the context of educational reporting are constructs which include a variety of information. Within this indicators different statistical based data are combined.



Indicators

Prof. Dr. Knut Schwippert

- **Indicators (in educational reporting)**
 - combine relevant information
 - can (!) be dynamic (depending on changing goals, new insights, and society)
 - cover a broad spectrum of (psychometric) indicators
- It follows that
 - indicators can change over time
 - indicators need to change over time



Indicators in Educational Reporting(s)

(e.g. HH 2009, BB 2008, Deu 2008)

Prof. Dr. Knut Schwippert

- Education in Germany
- Elementary Education, Care, and Education
- General Educational System and Non-Formal Learning Conditions
- Educational System of Apprenticeship
- University and High Schools
- Advanced Training and Education
- Transitions: School - Apprenticeship - University - Occupation
- Effects and Earning of Education

The Overarching Idea:
Education in the
Lifespan



Indicators

Prof. Dr. Knut Schwippert

Using indicators in the context of educational reporting seems to be the *agent of choice*

- it allows a detailed survey in the field of interest
- they are based on (objective) empirical bases
- historical and cultural conditions can be taken into account!



What will be observed in which context?

The perspective of
Comparative Education

Prof. Dr. Knut Schwippert



International and Comparative Research

Prof. Dr. Knut Schwippert

- With the perspective to combine ideas of international large scale surveys with comparative educational research it is important to take the historic perspective – which gives information about the origin and cultural imprint of an educational system – into account when results of the systematic empirical surveys are discussed.
- If one accepts historical and cultural differences between groups (e.g. countries, federal states, cities, districts) the results of comparative educational studies should be taken into account while systematic empirical studies are planned, conducted, and interpreted.

Hörner (1999)

Special Case: Germany



A Special Perspective: International Research

Prof. Dr. Knut Schwippert

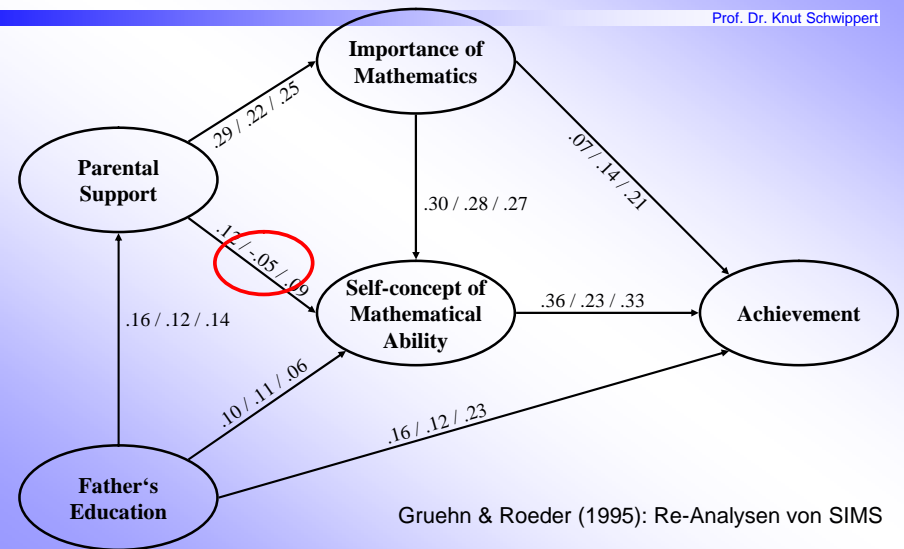
- Large-Scale Assessments: Empirical examination of characteristics and features of different countries and educational system.
- One key-question of this approach can be depicted as “what travels?”.
- For some research questions only international comparisons offer those variance components which are necessary to examine it’s effects within and between different units (like country or educational system).

Special Case: Germany



Path analysis of the relationship between motivational and parental background variables and student achievement in three countries (France / Japan / USA)

Prof. Dr. Knut Schwippert



The International Handbook of School Effectiveness Research

Based on SIMS

Table 8.3 Factors associated with effectiveness in different countries

Predictor Variable*	Country										
	Belgium (FL)	Belgium (Fr)	Canada (BC)	Canada (Ont)	Finland	France	Hong Kong	Hungary	Israel	Japan	Luxembourg
Father's occupation	15	16	22	25	39	40	43	44	50	54	59
Father's education	m	m		p	p	p		p	m		
Level of expected further education	p	p	p	p	p	p	p	m	p	p	p
Homework											
Teacher experience			n		p	p					p
Time spent keeping order					n	n					
Time spent on teaching	p						p				
Teacher expectations											
Use of published tests			p	p	p	p	p	p	p		p
Use of own tests							n	p			
Opportunity to learn	p	m		p	p	p	m		p	p	
Class size	n	p			p						
Urbanization								m	p		
Number of female teachers	p										n
Number of male teachers								p			
Number of meetings											

* Predictor Variables with significant Positive (p) or Negative (n) Associations with Mathematics Achievement.



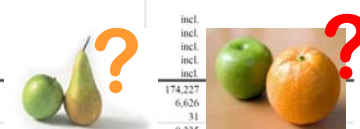
What Accounts for International Differences in Student Performance? A Re-Examination using PISA Data

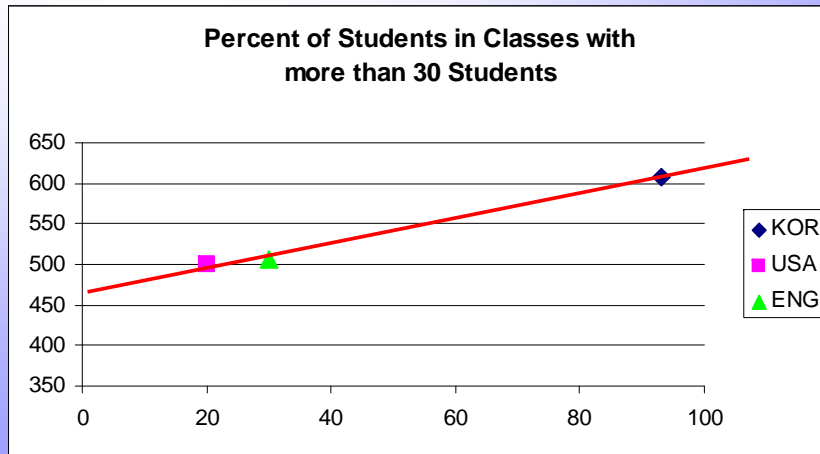
Prof. Dr. Knut Schwippert

	MATH			SCIENCE			READING					
	Coef	Std. Err.	Inter.	Std. Err.	Coef	Std. Err.	Inter.	Std. Err.	Coef	Std. Err.	Inter.	Std. Err.
Standardized tests	-6.381***	(2.389)	11.728***	(3.562)	-8.492***	(2.084)	15.836***	(3.333)	-6.306***	(2.159)	13.618***	(3.312)
School autonomy												
Determining course content	-6.197***	(2.473)	19.77***	(3.924)	-4.776***	(2.117)	14.272***	(3.929)	-8.622***	(2.275)	19.568***	(4.041)
Choosing textbooks	-2.017	(5.264)	48.295***	(10.106)	2.021	(4.054)	58.684***	(9.225)	0.092	(4.893)	61.049***	(10.074)
Formulating school budget	-4.176	(2.767)	-5.870**	(4.127)	-5.184***	(2.487)	-1.274	(3.920)	-0.121	(2.663)	-4.854	(4.106)
Deciding on budget allocations within school	5.495	(4.682)	11.649**	(6.933)	11.421***	(4.419)	5.588	(6.490)	7.491	(4.574)	4.281	(6.310)
Hiring teachers	16.503	(3.441)	-12.653**	(5.012)	2.719	(3.061)	-5.369	(4.988)	7.820**	(3.184)	-8.494**	(5.095)
Firing teachers	-4.197	(3.572)	4.791	(4.701)	-5.642*	(3.252)	6.564	(4.564)	-6.684**	(3.226)	2.749	(4.747)
Establishing teachers' starting salaries	-19.405***	(6.631)	23.332***	(7.956)	-5.118	(4.981)	3.398	(6.137)	-5.691	(3.599)	0.645	(6.810)
Determining teachers' salary increases	1.554	(6.388)	-1.931	(7.697)	-3.649	(4.692)	2.238	(3.899)	-1.639	(3.809)	0.699	(6.610)
Public vs. private operation and funding												
Publicly managed school	-16.677***	(3.518)	1.189	(4.878)	-12.572***	(2.654)	-0.419	(4.016)	-16.499***	(2.831)	5.267	(4.111)
Government funding (share)	8.119*	(4.909)	-0.252	(7.560)	-1.855	(3.967)	4.105	(6.509)	1.073	(4.053)	-1.724	(6.850)
Controls for:												
Student characteristics (8 variables)	incl.				incl.				incl.			
Family background (28 variables)	incl.				incl.				incl.			
Home incentives and inputs (6 variables)	incl.				incl.				incl.			
Resources and teachers (8 variables)	incl.				incl.				incl.			
Imputation dummies	incl.				incl.				incl.			
Students (units of observation)	96,855				96,758				174,227			
Schools (PSUs)	6,611				6,613				6,626			
Countries (strata)	31				31				31			
R ²	0.331				0.282				0.335			
R ² (without imputation controls)	0.310				0.255				0.312			

Notes: Dependent variable: PISA international test score. – Estimation models are the same as in Table 2, with one regression in each subject, only with the addition of interaction terms between external exit exams and other institutions. – Coef.: Coefficient estimate on the dummy (representing the effect in school systems without external exit exams). – Inter.: Coefficient estimate on the interaction term between the dummy and external exit exams (representing the difference in the effect between school systems without and with external exit exams). – Std. Err.: Clustering-robust standard error (taking account of correlated error terms within schools).

Significance level (based on clustering-robust standard errors): *** 1 percent, ** 5 percent, * 10 percent.





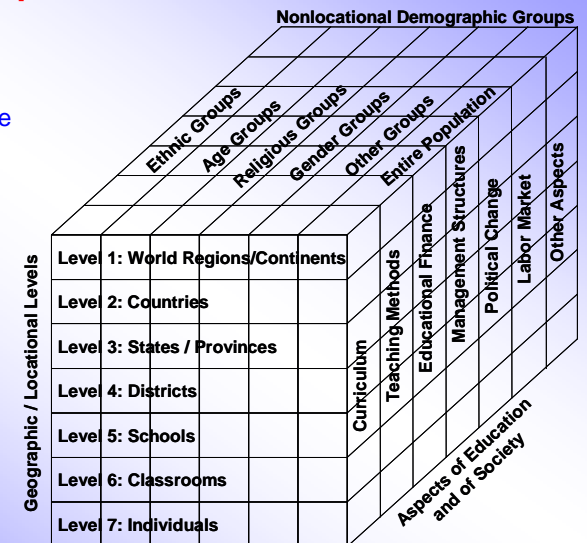
- If the average class size and the achievement of students are positive related:
Consequence 1:
Raising the class size
- If the teacher income and the achievement of students is negative related:
Consequence 2:
Reduce of teacher income

- If the average class size and the achievement of students are positive related:
Consequence 1:
Raising the class size
- If the teacher income and the achievement of students is negative related:
Consequence 2:
Reduce of teacher income

**Both consequences are wrong:
Both base on
misinterpretations due to the
ecological fallacy!**

Schematic model of a framework for comparative educational research

(Bray & Thomas, 1995)





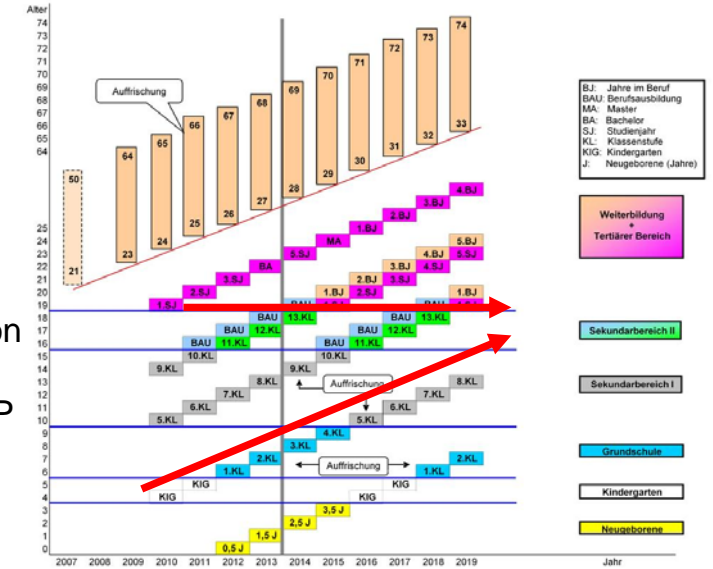
What will be observed and how long?

The question of time

Nobody can step into the same river twice

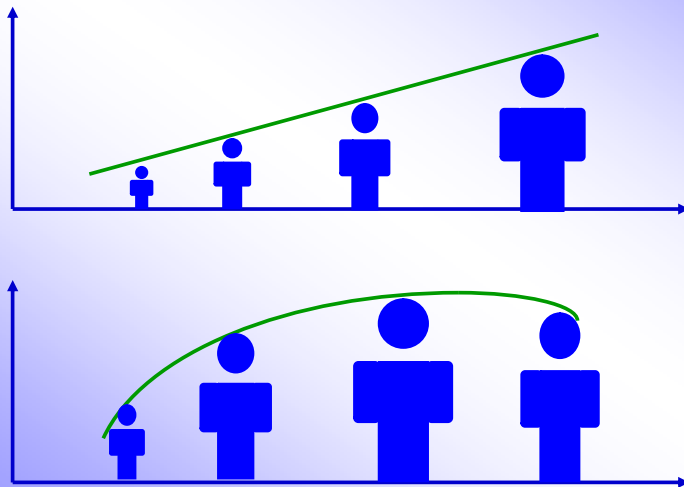


MULTI-KOHORTEN-SEQUENZ-DESIGN



• NEPS
Uni-Bamberg

• See the discussion in the DFG-SPP Kompetenz-Diagnostik



Future Perspectives



Implications for Research and Evaluation

Prof. Dr. Knut Schwippert

Six different types to test the model:

(Creemers & Kyriakides 2008: 213)

1. longitudinal studies on effectiveness
2. quantitative synthesis of effectiveness studies,
3. international studies on effectiveness, and
4. experimental studies.

5. Multilevel Analysis
6. Consideration of the expertise from comparative research

(Schwippert)



Some Trivial (?) Comments

Prof. Dr. Knut Schwippert

- What should be avoided in the context of Educational Reporting?
 - Simple (rough) comparisons
 - Lineup of studies with different scopes (argument against PISA-results „...but in my school...“)
 - For fear of innovations
- What must be learned in the context of Educational Reporting?
 - Training of empirical diagnostic (on different levels)
 - International comparisons to discover (traveling) innovations
 - Acceptance of inconvenient results



Future Perspectives

Prof. Dr. Knut Schwippert

This all works only if all parties understand each other

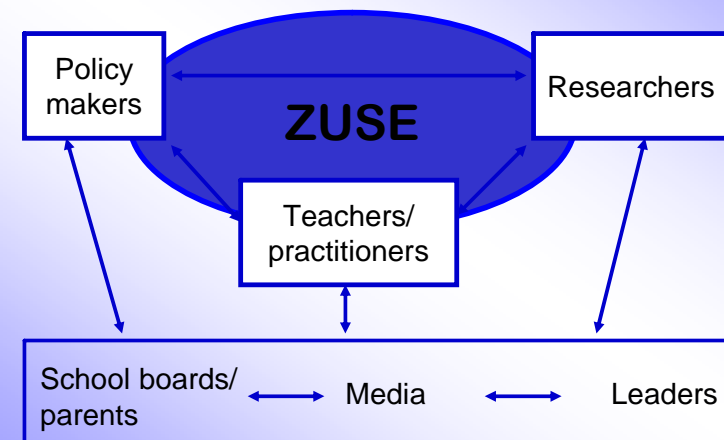
- It has to be accepted that policy-makers and researchers might have the same goals but do not have the same ways to reach these.

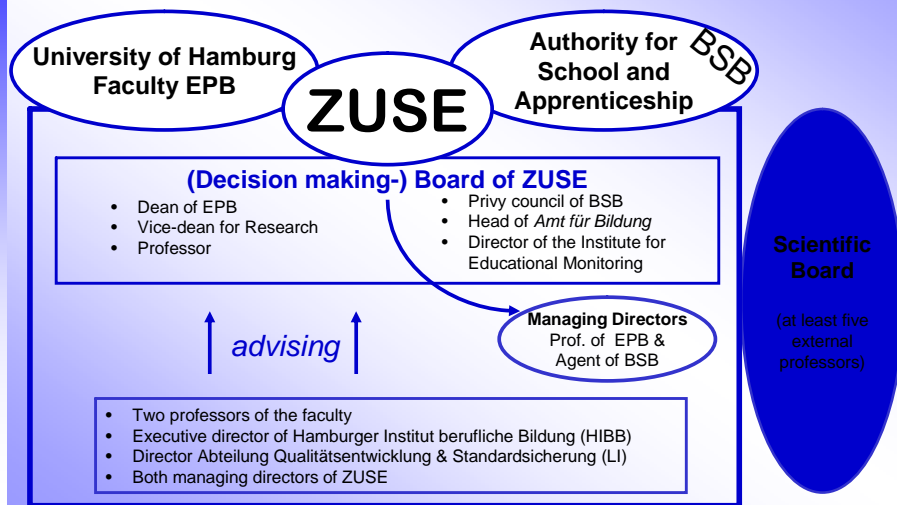
... coming back to the idea of *Brokerage Agencies*



Lines of Communication

Prof. Dr. Knut Schwippert





Many thanks for your patience !

Knut Schwippert

- Bray, M. & Thomas, R. M. (1995). Levels of Comparison in Educational Studies: Different Insights from Different Literatures and the Value of Multilevel Analysis. *Harvard Educational Review*, 65(3), 472-490.
- Centre for Educational Research and Innovation (CERI) (Hrsg.). (2007). *Evidence in education. Linking research and policy*. Paris: OECD.
- Creemers, B. P. M. & Kyriakides, L. (2008). *The dynamics of educational effectiveness. A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Gruehn, S. & Roeder, P. M. (1995). Concomitants of achievement in mathematics: A comparative analysis. In W. Bos & R. H. Lehmann (Hrsg.), *Reflections on educational achievements: Papers in honour of T. Neville Postlethwaite* (S. 88-106). Münster: Waxmann.
- Hörner, W. (1999). Historische und gegenwartsbezogene Vergleichsstudien - Konzeptionelle Probleme und politischer Nutzen angesichts der Internationalisierung der Erziehungswissenschaft. *Tertium Comparationis*, 5(2), 107-117.
- Husén, T. (1984). Introduction. In T. Husén & M. Kogan (Hrsg.), *Educational research and policy, How do they relate?* Oxford: Pergamon.
- Klieme, E., Baumert, J. & Schwippert, K. (2000). Schulbezogene Evaluation und Schulleistungsvergleiche. Eine Studie im Anschluss an TIMSS. In H.-G. Rolf, W. Bos, K. Klemm, H. Pfeiffer & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung (Band 11)* (S. 387-419). Weinheim: Juventa.
- Reynolds, D. (2000). School effectiveness: The international dimension. In H.-E. Ternoeth & D. Reynolds (Hrsg.), *The international handbook of school effectiveness research* (S. 232-256). London: Routledge.
- Ross, K. N. & Mählck, L. (1990). *Planning the quality of education. The collection and use of data for informed decision-making*. Oxford: Pergamon Press.
- Ternoeth, H.-E. (2004). Erfahrungen mit Beratung in Bildung und Wissenschaft: Bildungssystem und Bildungsforschung. In Deutsche Gesellschaft für Bildungsverwaltung (Hrsg.), *Bildungspolitik und Expertenmacht - Glanz und Elend der Politikberatung im Bildungswesen. Dokumentation de 24. DGBV-Jahrestagung vom 20. bis 22. November 2003 in Halle* (S. 51-64). Bonn: Deutsche Gesellschaft für Bildungsforschung.
- Woessmann, L. & Fuchs, T. (2004). What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data (Bd. Working Paper No. 1235): Center for Economic Studies & Ifo Institute for Economic Research (CESifo).