

Combining different types of data in educational research.

Data organization issues

T. Khavenson, Y. Tyumeneva

4th Conference of the European Survey Research Association,
Lausanne, Switzerland – July, 2011



National Research University – Higher School of Economics, Russia

What was our research initially about?

We studied trends of national achievements in PISA and TIMSS and their relation to some characteristics of the educational systems in post-soviet union and East-European countries (9 countries were selected).

Our **goal** was to find some effective ways of improving students' achievement, comparing PISA and TIMSS results with the trajectories of educational reforms and other socio-economic factors.

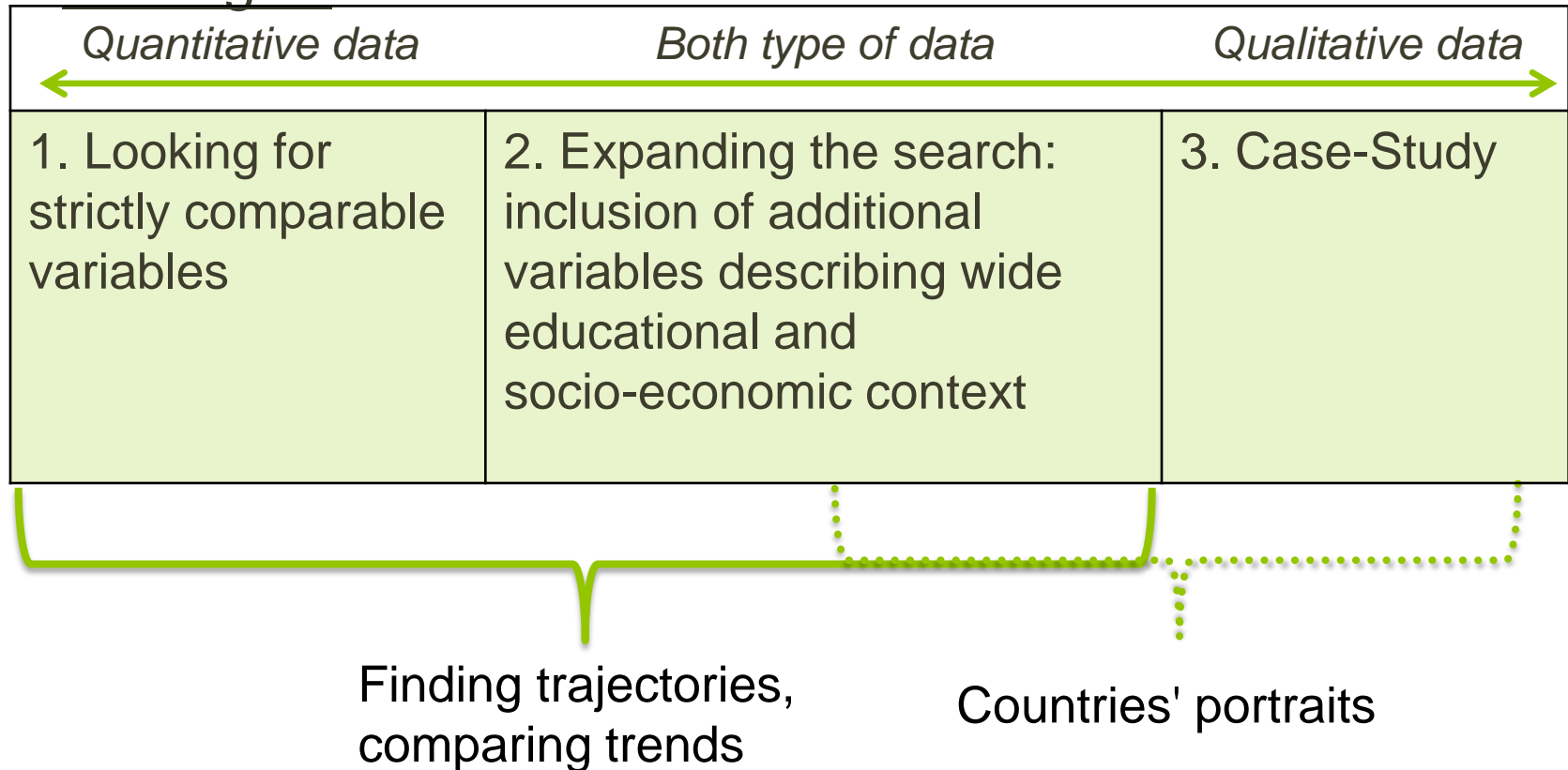
In addition it was important to evaluate the explanatory capacity of the data about the educational system (school level as well as country level) and the general socio-economic information too.

Data about achievements

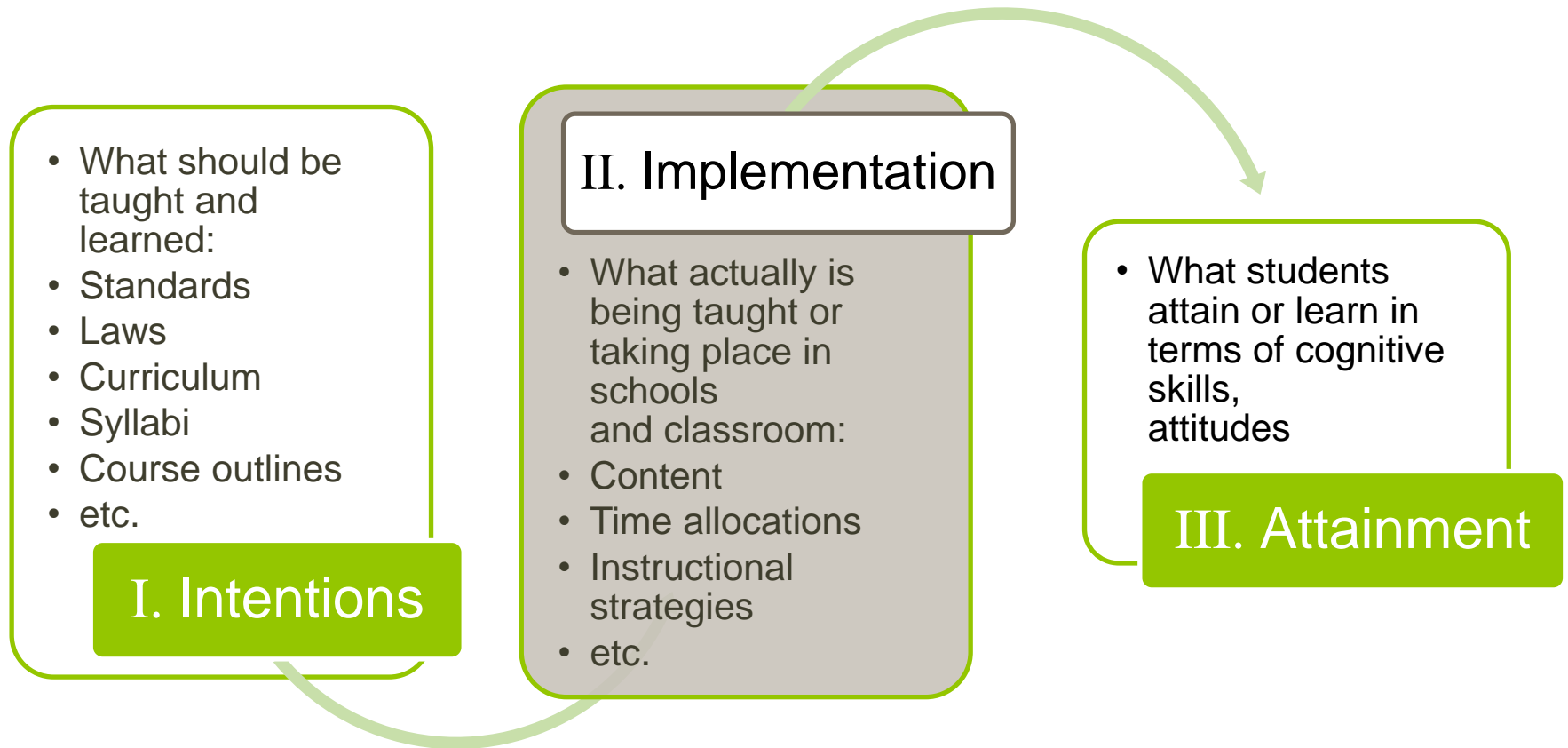
- ✓ TIMSS (Trends in International Mathematics and Science Study, IEA), 8th grade
- ✓ The data was from 4 waves: 1995, 1999, 2003, 2007
 - Math/science average scores
- ✓ PISA (Programme for International Student Assessment, OECD):
- ✓ The data was from 4 waves: 2000, 2003, 2006, 2009
 - Reading/math/science average scores
- ✓ Countries: Bulgaria, Czech, Hungary, Latvia, Lithuania, Russia, Romania, Slovakia, Slovenia

Finding information to help us explain

Strategies:



TIMSS and PISA data structure



In PISA terminology: Antecedents, Processes, Outcomes

Data type 1. Questionnaires

- ✓ Both PISA and TIMSS collect extensive background information about the context in which learning takes place. They provide data about school and classroom contexts, student characteristics and attitudes
- ✓ *"TIMSS makes every attempt to collect information about the important factors that foster improved teaching and learning in mathematics and science. The questionnaires concentrate on procedures and practices that have been shown to be effective in increasing achievement in mathematics and science. In this way, countries can better evaluate their TIMSS results." (TIMSS-2011: Assessment frameworks. Ch.3 Contextual Framework)*
- ✓ *"The primary role of the PISA conceptual structure for questionnaire development was to map the many components of existing models, to ensure that none of the essential dimensions are omitted from the data collection."*

Data type 1. Questionnaires

Variables and data sources

- Variables
 - Student achievements in math and science (TIMSS) and in math, science and reading (PISA)
 - Proficiency levels of students (distribution of countries' student performance)
 - School characteristics (location, resources)
 - Student's educational practices
 - Teacher's classroom practices
- Data sources
 - School principals, teachers and students questionnaires
TIMSS and PISA (aggregated at the country level)

Data type 1. Questionnaires

Benefits and restrictions

- Benefits
 - Comparability
 - Representative sample
 - Wide range of issues discussed
- Restrictions
 - Items changing from year to year
 - Social desirability (especially in teacher's answers)
 - Inter-rater agreement (Fleiss Kappa coefficient) – average value of this coefficient is 0.11. As for good Inter-rater agreement we need 0.6 as minimum (Landis and Koch 1977).

Data type 1. Questionnaires

Restrictions example

The main problems:

- Response options mismatch
- Omitting items in some waves

TIMSS 2007, 2003	TIMSS 1999, 1995
How many people live in the city, town, or area where your school is located?	In what type of community is your school located?
More than 500,000 people	A geographically isolated area
100,001 to 500,000 people	Village or rural (farm) area
50,001 to 100,000 people	One on the outskirts of a town/city

- ✓ Other examples are: rules of school admission, school resources, school and class supervision, responsibility for different aspects of school life etc.
- ✓ In our research we could only choose a short list of school resources from this type of data

Data type 2. Country encyclopedias

The information used in the TIMSS 2007 Encyclopedia is intended to complement the quantitative data that is contained in the International Mathematics and Science reports.

- Variables (only quantitative data):
 - Intended curriculum
 - Countries' education systems and policies (pupil/teacher ratio, prerequisites to be a teacher etc.)
 - Precise information about educational reforms (national standard, main years and events of reforms etc.)

- Data source: Countries' encyclopedias in TIMSS and PISA

Data type 2. Country encyclopedias.

Benefits and restrictions

- Benefits
 - Various information
 - Countries provide information themselves
- Restrictions
 - Information about intentions mostly
 - Countries may only provide the information they want to share
 - There is quite a lot of unique information that is not comparable with other countries

In particular, it concerns information about curricula, national educational standards and the process of educational reforms

Data type 3. Socio-economic data

Selected variables were based on the theoretical assumptions about their corresponding with educational outcomes

- Variables (only quantitative data):
 - Enrolment in primary, secondary education
 - Expenditure per student (% of GDP per capita) in primary, secondary, tertiary education
 - Public spending on education
 - Literacy rate
 - GDP per capita
 - Urban population
 - High-technology exports
 - etc.
- Data source:
 - <http://data.worldbank.org> The World Bank Open Data
 - <http://stats.uis.unesco.org> UNESCO Institute for Statistics
 - www.eurydice.org Eurydice - Information on Education Systems and Policies in Europe (European Commission)

Data type 3. Socio-economic data.

Benefits and restrictions

- Benefits
 - Wide range of issues
 - National characteristics
- Restrictions
 - Lack of certain data for selected countries.
 - Lack of comparable data

Data type 4. International publications

- International publications
 - Articles, reports, conference papers and so on where authors analyzed their national educational systems and reforms
- Data source:
 - Publications in academic journals
 - Eurydice - Information on Education Systems and Policies in Europe (European Commission) www.eurydice.org
 - Countries' encyclopedias in TIMSS and PISA
- *No specific variables for this type of data were identified*
- Benefits
 - Deep and comprehensive analysis
- Restrictions
 - Articles intend to focus on issues that are "on trend" and don't always imply the real biggest or the most important problems
 - Few relevant sources
 - Almost all reports and a lot of the articles are written in the national languages

Conclusions. Explanatory capacity

Data type	Explanatory capacity (expected)	Explanatory capacity (observed)	What did we get
Questionnaire	Implementation	Intentions	Resources trends, achievements trends, school SES
Socio-economic data	Implementation Attainment	Implementation Attainment With omissions and less comparable than it is necessary	Partly unique diagrams with trajectories of all included variables.
Country encyclopedias	Intentions Implementation Attainment	Intentions Only formal implementation	
International publications	Intentions Implementation Attainment	Implementation Attainment	Unique country's portraits (the reform process)

Findings and conclusion

- ✓ Portraits of countries. The deeper the portrait one wants, the less comparable the data will be. It is desirable to have an expert evaluation of each country to deepen the portraits
- ✓ School level information quite often subjects to social desirability and low inter-rater agreement. So one should mostly use factual information rather than self-appraisal
- ✓ Descriptive information from countries' encyclopedias is usually about intentions. So it can't be used when the research question concerns implementation or attainment. Or it is necessary to carry out additional research that can provide standardized and comparable data

Discussion and recommendations

- ✓ Data combining strategy gives useful results for each specific country, however it is less useful in comparative research
- ✓ There is a lack of information on social and economic context (socio-economic data and country encyclopedias), especially concerning non-European countries. So it makes it difficult to carry out comparative research using this information
- ✓ The TIMSS 2011 contextual framework includes information about National and Community Contexts. Therefore it will provide fully comparable information

Literature

- ✓ Mullis, I., Martin, M., Ruddock, G., O'Sullivan, C., Preuschoff, C. (2009)). TIMSS-2011: Assessment frameworks. TIMSS&PIRLS International Study Center Lynch School of Education, Boston College, p. 93.
- ✓ OECD (2009). Technical Report for the OECD Programme for International Student Assessment 2006, OECD Publishing, p. 51
- ✓ Plomp, T. (1998). The potential of international comparative studies to monitor the quality of education. Prospects, 28, 45-59

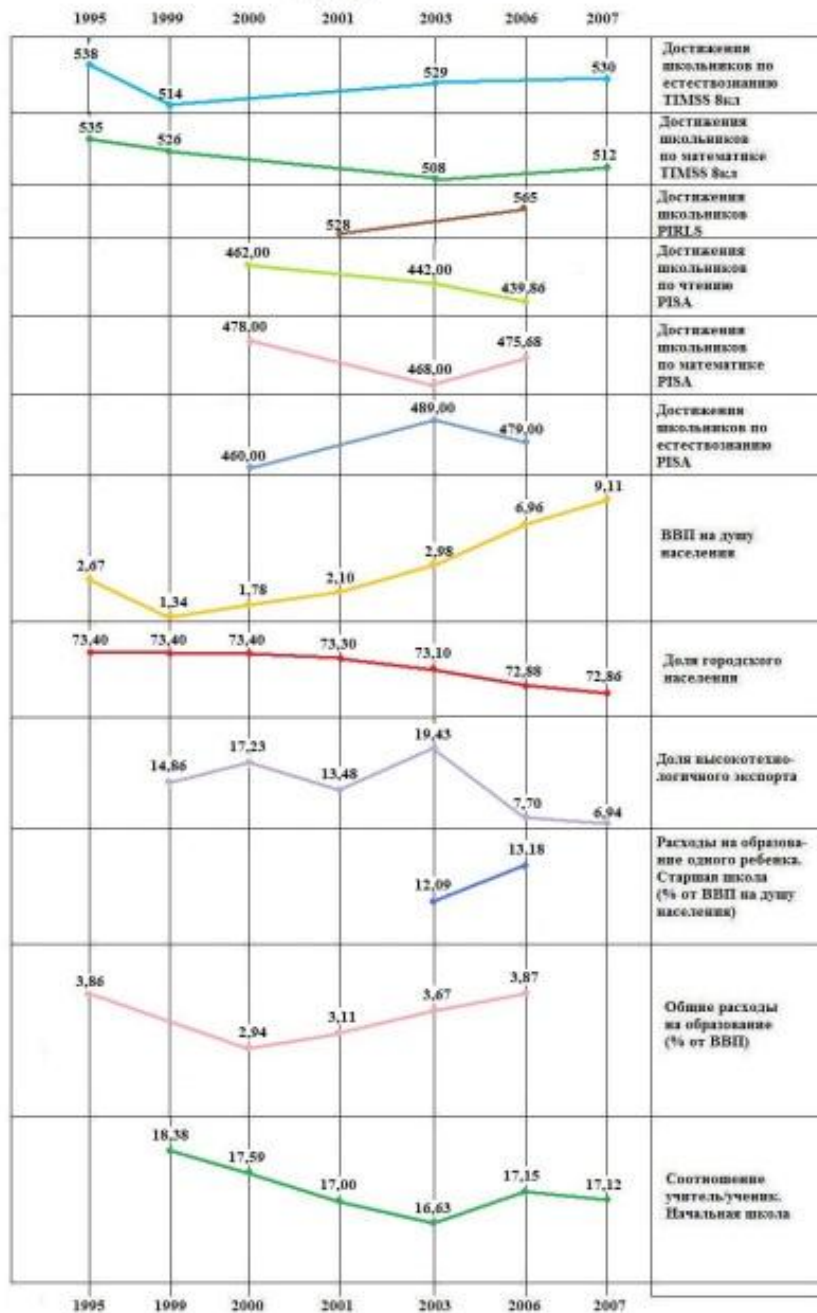
***Thank you for your
attention!***

We used presentation templates from website www.presentationmagazine.com

27.07.2011

T.Khavenson, Y.Tyumeneva. HSE,
Moscow

Россия



27.07.2011