

# DDI for Large Scale Assessments in Education

## *Data Archiving and Question Banking*

**Martin Mechtel & Michel Knigge**

***EDDI, First Meeting, IZA Bonn***

***December 4<sup>th</sup>, 2009***

### **Overview**



Institut zur Qualitätsentwicklung  
im Bildungswesen  
Forschungsdatenzentrum

- Development of educational large scale assessments in Germany
- Institute for Educational Progress (IQB)
- Special properties of test items in educational assessment
- Special properties of sample designs in educational large scale assessments
- Special properties of scaling in educational assessment
- Documentation “standards” at present and the EduDDI network
- Summary

## **Germany was in empirical educational research in the 60s ...**

- and participated in occasional international Large Scale Assessments (e. g. FIMS, 1964)

## **In the 70s and 80s Germany retreated from large scale educational research and ...**

- did not participate in international Large Scale Assessments anymore (e. g. SIMS, 1978)
- neglected the development of Programs for Educational Monitoring
- neglected empirical Research in Educational Science

→ one important reason was the very controversial debate about comprehensive schools

## **The Empirical Turnover in Germany started in the 90s ...**

- with the setup of and participation in several studies (e. g. BIJU, 1990; TIMSS, 1995)

## **At the beginning of the new millenium...**

- the PISA-2000-shock hit Germany
- as a reaction Educational Monitoring was widened up: E. g. PIRLS, LAU, MARKUS, but especially the setting of Educational Standards (Output Orientation) → IQB

## **The Institute for Educational Progress**

- Founded in 2004 to further develop and measure the educational standards of the Standing Conference of the Ministries of Education
- Test development: Assess student's competencies in certain subjects at certain stages of the educational system
- Funded by the Standing Conference of the Ministries of Education
- Affiliated to the Humboldt-University Berlin
- Regular funds about € 4.5 Mio/year,
- Staff about 50 regular plus 50 students
- Since 2007: Research Data Center (FDZ), financed by Federal German Ministry of Education and Science; staff 4 plus 2

**Vogonische Lyrik**

Grupp, ich beschwöre dich, mein punzig Turteldrom  
 Und drängel reifig mich mit kranken Bindelwördelein  
 Denn sonst werd ich dich rändern in deinen  
 Gobberwarzen  
 Mit meinem Börgelkranze, wart's nur ab!

**Aufgabe 1**

Welches Sprichwort passt am ehesten zu der Aussage des Gedichts?  
 Kreuze die richtige Antwort an.

Reisende soll man nicht aufhalten  
 Liebe macht blind  
 Ohne Fleiß keinen Preis  
 Nachts sind alle Katzen grau.

An welcher Zeile ist das erkennbar (Zeilennummer):  
 \_\_\_\_\_

**Aufgabe 2**

Nenne zwei Klangfiguren in diesem Gedicht:  
 \_\_\_\_\_  
 \_\_\_\_\_

**Task M042**

**Stimulus**

**Subtask A**

**Item 01**  
**Variable a**

**Item 02**  
**Variable a**

**Subtask B**

**Item 03**  
**Variable a**  
**Variable b**

```

<QuestionScheme id="M042"
    urn="urn:ddi:3_0:QuestionScheme=
        igh.de.ddi:MATH2008D_M042(1_12)">
    <MultipleQuestionItem id="M042A">
        <Group id="M04201">
            <QuestionItem id="M04201a">
                ...
            </QuestionItem>
        </Group>
        <Group id="M04202">
            <QuestionItem id="M04202a">
                ...
            </QuestionItem>
        </Group>
    </MultipleQuestionItem>
    <MultipleQuestionItem id="M042B">
        <Group id="M04203">
            <QuestionItem id="M04203a">
                ...
            </QuestionItem>
            <QuestionItem id="M04203b">
                ...
            </QuestionItem>
        </Group>
    </MultipleQuestionItem>
</QuestionScheme>

```

## Special Properties of Sample Designs

test design (example)		booklets														
timeline		1	2	3	4	5	6	7	8	9	10	11	12	13	UH	
Cluster		1	M1	M2	M3	M4	M5	M6	M7	S1	S2	R1	R2	PS1	PS2	PS/R
60 Min	1	M2	M3	M4	M5	M6	M7	S1	S2	R1	R2	PS1	PS2	M1	M/S	
60 Min	2	10 Min	Pause													
60 Min	3	M4	M5	M6	M7	S1	S2	R1	R2	PS1	PS2	M1	M2	M3		
	4	15 Min	R1	R2	PS1	PS2	M1	M2	M3	M4	M5	M6	M7	S1	S2	
47 Min		Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q		
total 167 Min																
Legend:		M = Mathematics; S = Science; R = Reading; PS = Problemsolving; Q = Questionnaire; UH = Booklet for students with special education needs ("Une heure")														

- Before Scaling: Coding and Scoring

### **Example: Capital of Germany?**

**Answer – Code – Score**

**Berlin – 101 – 1**

**berlin – 102 – 1**

**Bonn – 111 – 0**

...

- Scaling Model needs to be documented for each Scale

### **Background Scales**

**Computation of Mean / Sum?**

**Missing Values / Imputation?**

### **Assessment**

**Background Model (SES, Gender, Migration, ...)**

**Weights**

**Dimensions (1 competency= 1 dimension / simultaneously?)**

**Parameters (classical: Rasch-Model = 1 parameter difficulty, but possible up to three)**

## **Documentation “standards” at present and the EduDDI network**

- Improvement of the data documentation is necessary
  - Currently very low standards in data documentation in educational research
  - Usually only codebooks, some studies have a technical report
  - What is documented and how is somehow arbitrary (inhouse solutions)
- Starting up the EduDDI work group
  - Implement the DDI 3.0 standard in educational research
  - Cooperation with central educational research institutions like NEPS, IEA
  - Objectives
    - Easier Access (archiving, re- and secondary use)
    - Synergies between institutions (item and knowledge exchange)
    - Improvement of internal workflows (e.g. production of test booklets, scaling)

- Since the Empirical Turnover Educational Assessments are well established in Germany
- Through a higher level of metadata standardization we hope for improvements in synergies between institutions, internal workflows, and archiving
- DDI 3 provides a framework that can be used for this purpose
- Still extra standardization for the special requirements of Educational Assessment Data is necessary
- Especially the full tasks with the interdependency of the items, the multi-matrix sample designs, and the scaling models have to be standardized
- The foundation of the EduDDI network in 2008 was an important step on the way to accomplish this task

---

Martin Mechtel & Michel Knigge – DDI for Large Scale Assessments in Education – Question Banking and Archiving

9



*Thank you!*

