



Computer-based Formative Assessment in Classrooms

Stéphanie Berger¹², Urs Moser¹, Angela Verschoor³ & Theo Eggen²³

1 Institute for Educational Evaluation, Associated Institute of the University of Zurich

2 Research Center for Examinations and Certification, University of Twente

3 CITO

AEA-Europe Conference, Limassol, 3th to 5th November 2015



Outline

- Introduction
 - Educational assessment in Northwestern Switzerland
 - Formative assessment
- Methodological Approach
 - Competence-based curriculum
 - Item Response Theory
- Implemented Assessment Types
 - Adaptive assessments
 - Mastery assessments
 - Diagnostic assessments
- Discussion



**Universität
Zürich** ^{UZH}

Institute for Educational Evaluation
Associated Institute of the University of Zurich



INTRODUCTION



Educational Assessment in Northwestern Switzerland

- Initiative of four cantons in Northwestern Switzerland
- N = 13'000 students per school year
- Instruments:
 - Four compulsory standardized tests
 - Online item bank for formative assessment
- Assessment of student ability in five different school subjects
 - German (first language)
 - English, French
 - Mathematics
 - Science





Formative Assessment

Van der Kleij, Vermeulen, Schildkamp & Eggen, 2015

- Assessment to support learning and to accommodate students' individual educational needs
- Focus on student and class level
- Three different approaches
 - Data-based decision making
 - Assessment for Learning
 - Diagnostic testing



Data-based Decision Making (DBDM)

Schildkamp & Kuiper, 2010; Van der Kleij, Vermeulen, Schildkamp & Eggen, 2015

- Focus on learning outcome
- Systematic collection and analysis of objective data
- Feedback loop
 - Assessment of students' current ability level
 - Definition of appropriate learning goals based on assessment outcomes
 - Monitoring/evaluation of students' progress through further assessments
 - Adjustment of goals and/or learning environment based on the assessment outcomes



Online Item Bank www.mindsteps.ch



The screenshot shows the homepage of the Mindsteps website. At the top, there is a browser address bar with the URL <https://www.mindsteps.ch/de/>. Below the address bar is the Mindsteps logo, which consists of a stylized 'M' in green and blue followed by the text 'MINDSTEPS'. To the right of the logo are navigation links: 'Konzept', 'FAQ', 'KONTAKT' (in an orange button), and 'ANMELDEN' (in a blue button). The main content area features a large photograph of three children in a classroom. The foreground shows a young girl with brown hair, smiling broadly. In the background, two other children are visible, looking down at their work. Overlaid on the image is the text 'Lernen sichtbar machen' in a large, white, sans-serif font. Below this, in a smaller white font, is the text 'Mindsteps ist eine Aufgabensammlung für die Unterstützung des kompetenzorientierten Lernens.' In the top right corner of the image area, there is an orange circular badge with the text 'Version 1.0'.



Online Item Bank www.mindsteps.ch

- Accessible through a web-based platform
- Item bank with several thousands of items
- Three assessment types measuring on different levels
- Assignment of assessments to selected students
- Computer-based (adaptive) assessment delivery
- Automated scoring
- Immediate reporting



**Universität
Zürich** ^{UZH}

Institute for Educational Evaluation
Associated Institute of the University of Zurich



METHODOLOGICAL APPROACH



Competence-based Curriculum

(www.lehrplan.ch)

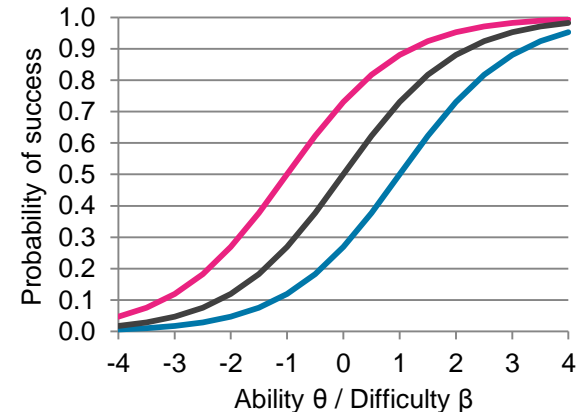
- Hierarchical structure
 - Subject
 - Domain
 - Competences
 - Competence levels
- Content framework
 - Item development
 - Item selection
 - Criterion-referenced feedback

MA.1	Zahl und Variable
A	Operieren und Benennen
3. Die Schülerinnen und Schüler können addieren, subtrahieren, multiplizieren, dividieren und potenzieren.	
MA.1.A.3 Die Schülerinnen und Schüler ...	
1	a » können im Zahlenraum bis 20 ohne Zählen verdoppeln, halbieren, addieren und subtrahieren.
	b » können bis 100 ohne 10er-Überträge addieren und subtrahieren ohne Zahlen (z.B. $35 + 13$) » können auf den nächsten 10er ergänzen. » können bis 100 verdoppeln (5er- und 10er-Zahlen) und halbieren (10er-Zahlen). » können zweistellige Zahlen in 10er und 1er zerlegen (z.B. 25 in zwei 10er und fünf 1er).
	c » können im Zahlenraum bis 100 verdoppeln, halbieren, addieren und subtrahieren. » kennen Produkte aus dem kleinen Einmaleins mit den Faktoren 2, 5 und 10. » können Produkte aus dem kleinen Einmaleins in Faktoren zerlegen (z.B. $36 = 6 \cdot 6 = 4 \cdot 9$).
2	d » können beim Addieren und Subtrahieren Rechenwege notieren und Ergebnisse überprüfen. » können schriftlich addieren und subtrahieren. » kennen die Produkte des kleinen Einmaleins.
	e » können bis 4 Wertziffern im Kopf addieren und subtrahieren (z.B. $320'000 + 38'000$; $402 + 90$). » können bis 4 Wertziffern multiplizieren (im Kopf oder mit Notieren eigener Rechenwege, z.B. $45 \cdot 240$). » können natürliche Zahlen durch einstellige Divisoren dividieren (im Kopf oder mit Notieren eigener Rechenwege, z.B. $231 : 7$).
	f » können Dezimalzahlen bis 5 Wertziffern addieren und subtrahieren (im Kopf oder mit Notieren eigener Rechenwege, z.B. $30.8 + 5.6$). » können Brüche mit den Nennern 2, 3, 4, 5, 6, 8, 10, 20, 50, 100 am Rechteckmodell kürzen, erweitern, addieren und subtrahieren. » können Grundoperationen mit dem Rechner ausführen.
3	g » können Dezimalzahlen bis 5 Wertziffern multiplizieren und die Ergebnisse überprüfen (im Kopf oder mit Notieren eigener Rechenwege, z.B. $308 \cdot 52$; $12 \cdot 0.3$). » können Brüche mit den Nennern 2, 3, 4, 5, 6, 8, 10, 20, 50, 100 am Rechteckmodell multiplizieren. » können Brüche mit den Nennern 2, 3, 4, 5, 6, 8, 10, 20, 50, 100, 1'000 als Dezimalzahlen schreiben. » können bestimmen, wie oft Stammbrüche in ganzen Zahlen enthalten sind (z.B. Wie viele Male ist $\frac{1}{6}$ in 2 enthalten? $\rightarrow 2 : \frac{1}{6}$).
	h » können Prozentrechnungen mit dem Rechner ausführen. » Erweiterung: können natürliche Zahlen in Primfaktoren zerlegen.

Item Response Theory (IRT)

- Measuring item difficulty and student ability on a single scale
- Rasch Model (Rasch, 1960)

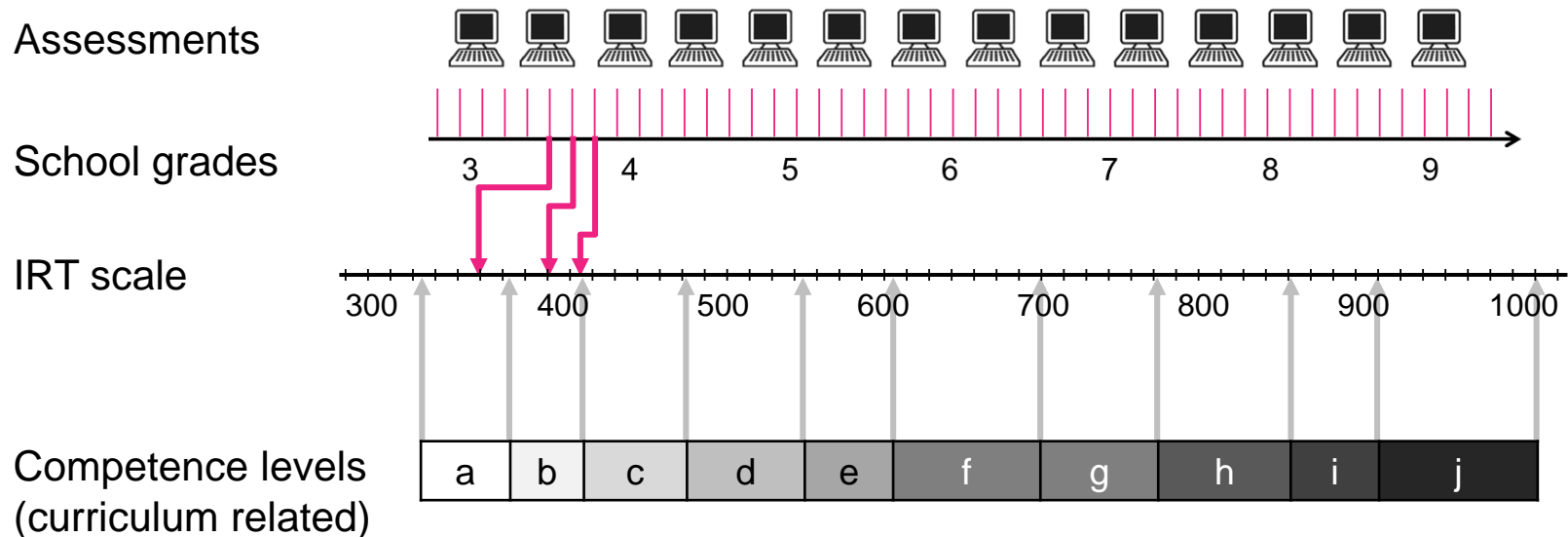
$$P(X_{ij} = 1 | \theta_i, \beta_j) = \frac{\exp(\theta_i - \beta_j)}{1 + \exp(\theta_i - \beta_j)}$$



- Precondition: Calibrated item bank
 - Collect empirical information in pretests
 - Automated online calibration algorithm

IRT Scale for Item Selection and Reporting

- Criterion-referenced interpretation → meaningful scores
- Compare results over time → progress reporting





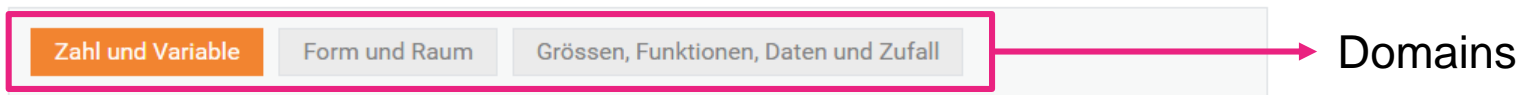
**Universität
Zürich** ^{UZH}

Institute for Educational Evaluation
Associated Institute of the University of Zurich



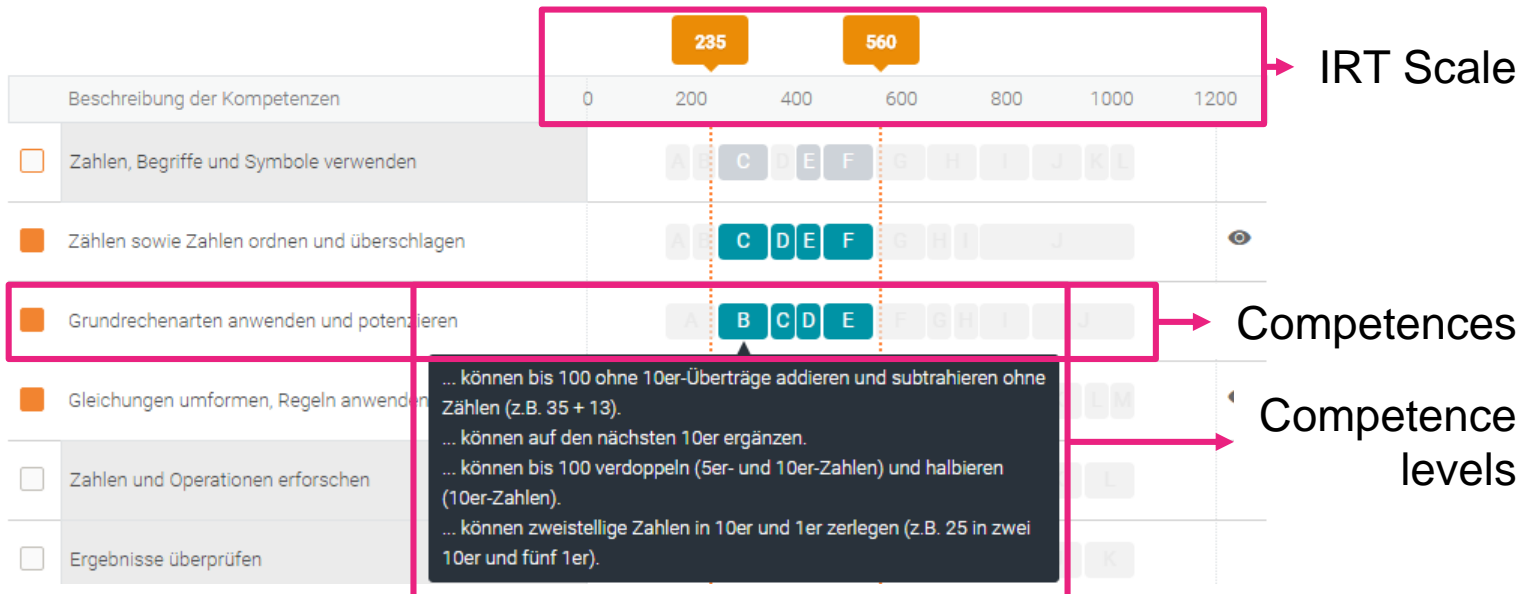
IMPLEMENTED ASSESSMENT TYPES

Implementation of the Item Bank



Für Ihre aktuelle Auswahl stehen **800** Aufgaben zur Verfügung.

Nicht verfügbar Verfügbar Ausgewählt



Adaptive Assessments – Feedback on General Ability Level

- Assessment specification:
 - Assessment scale (i.e., subject/domain)
 - Starting difficulty
 - Target difficulty (cf. Eggen & Verschoor, 2006)
- Test administration: adaptive
- Reporting:
 - Scale level (i.e., subject or domain)
 - Progress reports

Fach- und Kompetenzbereiche *

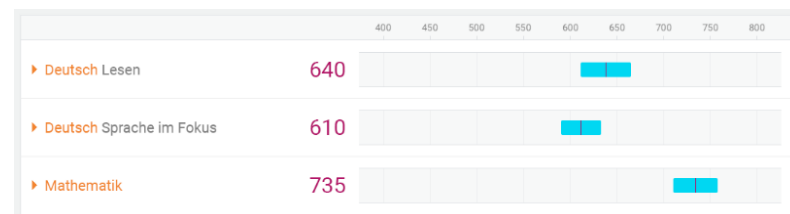
Mathematik Demo ▾

Startschwierigkeit *

670

Testschwierigkeit *

mittel ▾



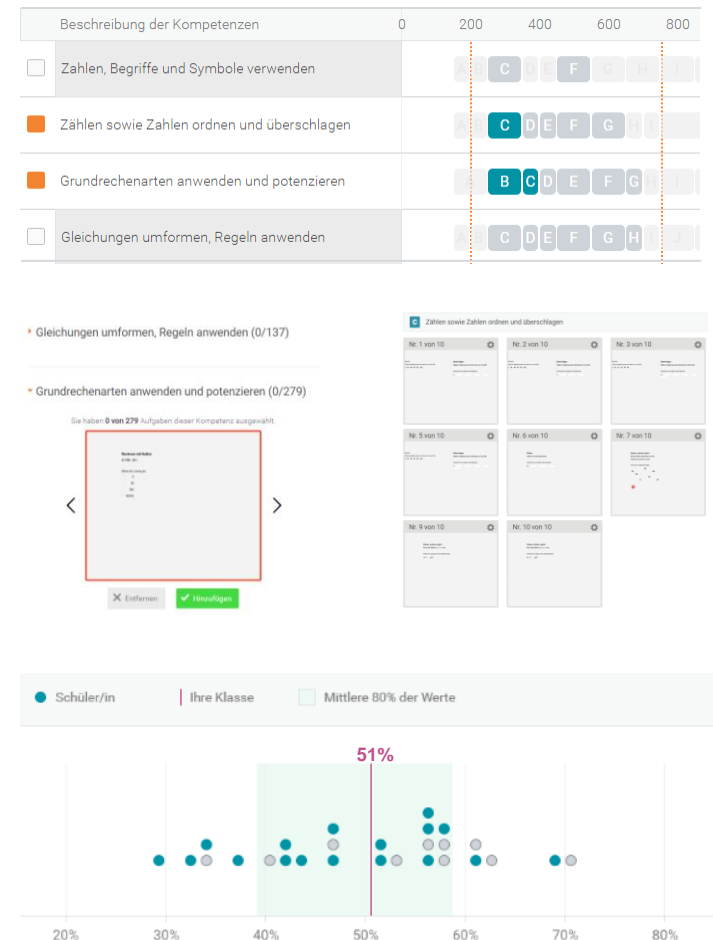
Mastery Assessment – Feedback on Competence Level

- Assessment specification:
 - Assessment scale (i.e., subject/domain)
 - 1-3 competence levels
 - Item review through teacher
- Test administration: linear
- Reporting:
 - Competence level
 - Percentage mastery
 - Identification of strengths and weaknesses



Diagnostic Assessment – Feedback on Assessment Level

- Assessment specification:
 - Assessment scale
 - Items
 - Assessment length
 - Item order
- Test administration: linear
- Reporting:
 - Assessment level
 - Item level
 - Similar to traditional class assessments





**Universität
Zürich** ^{UZH}

Institute for Educational Evaluation
Associated Institute of the University of Zurich



DISCUSSION



Discussion

- Computer-based assessment instrument for assisting teachers in formative assessment (i.e., data based decision making)
- Item bank structured based on the New Swiss Curriculum
- IRT as a powerful measurement approach for
 - Selecting suitable items,
 - providing competence-based feedback,
 - and for comparing assessment results over time
- Complementary to already existing teaching material



THANK YOU FOR YOUR ATTENTION!



Contact: Stephanie.Berger@ibe.uzh.ch



References I

- Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice* 18(1), 5–25.
- Berger, S., Verschoor, A. J., Moser, U., & Oostlander, J. (2014). *Educational Assessment in Northwestern Switzerland: Psychometrical challenges in linking standardized tests to an online item bank for formative assessment*. Paper presented at the AEA-Europe Conference, Tallinn, Estonia.
- Deutschschweizer Erziehungsdirektoren-Konferenz (D-EDK). (2015). *Lehrplan 21*. Retrieved from <http://www.lehrplan.ch>
- Eggen, T. J. H. M., & Verschoor, A. J. (2006). Optimal Testing With Easy or Difficult Items in Computerized Adaptive Testing. *Applied Psychological Measurement*, 30, 379–393.
- Hattie, J. A. C., & Brown, G. T. L. (2007). Technology for school-based assessment and assessment for learning: Development principles from New Zealand. *Journal of Educational Technology Systems*, 36(2), 189–201.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and Practices* (Third Edition). New York Heidelberg Dordrecht London: Springer.
- Moser, U. (2009). Test. In S. Andresen, R. Casale, T. Gabriel, R. Horlacher, S. Larcher Klee, & J. Oelkers (Eds.), *Handwörterbuch Erziehungswissenschaft* (pp. 866–880). Weinheim und Basel: Beltz.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.



References II

- Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, 26, 482–496.
- Van der Kleij, F. M., Vermeulen, J. A., Schildkamp, K., & Eggen, T.J.H.M. (2015). Integrating data-based decision making, Assessment for Learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice* (ahead-of-print), 1–20.
- van der Linden, W. J., & Glas, C. A. W. (2010). *Elements of adaptive testing*. New York: Springer.
- Wainer, H. (2000). Introduction and history. In H. Wainer (Ed.), *Computerized adaptive testing. A primer* (2nd ed., pp. 1–21). Mahwah NJ.: Lawrence Erlbaum Associates.
- Wainer, H., & Mislevy, R. J. (2000). Item response theory, item calibration, and proficiency estimation. In H. Wainer (Ed.), *Computerized adaptive testing. A primer* (2nd ed., pp. 61–100). Mahwah NJ.: Lawrence Erlbaum Associates.