NORDIC
METRICS
Danish experiences with It-based assessment
Jakob Wandall,
NordicMetrics, partner - director R&D

NORDIC METRICS

Background

→ 1994 Research, social science

→ 2011 Ministry of Education, 15yrs.

2011 → Consultant, edu. testing/assessment

2013 → NordicMetrics, partner, director R&D

NORDICMETRICS

NordicMetrics arbejder for at forbedre kommuners, skolers og læreres grundlag for den pæda-gogiske indsats og for at give eleverne adgang til inspirerende læringsmaterialer hvor træning kombineres med konstruktivt feedback.

Vi arbejder mest med analyser, måling af og træning i læsning og fremmedsprog, men arbej-der også med andre fagområder og felter (fx det naturvidenskablige område, måling af trivsel, selvtillid og motivation og andre personlige egenskaber, QA på test og

NORDIC METRICS

Content

1. Background for Danish School Policy
2. MEX vs. DK – Culture and Values
3. Formal Assessment in DK in the Basic school
   A. Traditional exams
   B. The National Testing system
4. Lessons Learned

NORDIC METRICS

Facts about Denmark

Denmark

- 4,9 mill. Inhabitants
- 90% ethnic Danes

The public primary and lower secondary education in Denmark, ”Folkeskolen”:

- 600.000 students, 60.000 pr. Grade level
- App 2.000 public schools, of those 1.300 ordinary folkeskoler,
- School governed by decentralized local governments → significant differences from municipality to municipality
- 3-5 students pr. new computer (<3 year) – some place 1:1
- highspeed internet-connection almost everywhere

→ **Infrastructure suited for It-based assessment**

Formal assessment in Denmark, ”Folkeskolen”:

- Summative – Final examination after 9th grade
- Formative – National testing system (10 adaptive test from 2nd-8th grade)

NORDIC METRICS

Background for the Danish school policy

- **1814 – Protestant church & The nation state: First dedicated School law.**
  - **Double purpose:**
    - To bring up the children to upright, law-abiding, hard-working **Christian** human beings
    - Provide Skills & knowledge → Useful citizens for the nation state (national army)
  - Focus on minimum requirements: Necessary skills & knowledges
- **1970 - Reform of governance: Decentralization.**
  - Schools (economy, curriculum, teachers) were decentralized from state to 275 local governments
  - Only legislative framework remained on state level
  - Alliance between teacher organization & state
- **1993 - Globalization: EU and International surveys**
  - IEA 1990 - First participation in international surveys
  - Introduction of Evaluation in the school regulation
  - Policy focus ***from Teaching to Learning*** – low/no impact in the schools
  - Alliance between teacher organization & state crumbles
- **2006 – New concept: Culture of evaluation**
  - National testing system
  - Final exams made mandatory

NORDIC METRICS

The Danish Act on public primary and lower secondary school (2014)

Section 1. **STATEMENT OF PURPOSE:**

(1) The school must - in cooperation with the parents – give the students **knowledge and skills** that:

1. prepare them for further education and **make them want to learn more**,
2. make them familiar with Danish culture and history,
3. give them an understanding of other countries and cultures,
4. contribute to their understanding for human interaction with nature and
5. promotes the individual student's **personal overall development**.

(2) The school must develop working methods and provide a framework for experience, reflection and dynamism so students **develop awareness and imagination and confidence in own ability, makes a stand and take action.**

(3) The school should **prepare students for participation, joint responsibility, rights and duties in a society of freedom and democracy**. School work must therefore be characterized by intellectual freedom, equality and democracy.

The "Double purpose " of 1814 still very clear:

1. Cognitive development (knowledge & skills) are defined ***as means***
2. The overall objectives are non-cognitive: Grid, motivation, efficacy , attitudes and values (Christianity → Democracy).

Comparing
Denmark
&
Mexico

NORDIC METRICS

World Value Surveys (an Cross-cultural value study)

Wave 1-6, Surveys 1981-2014:

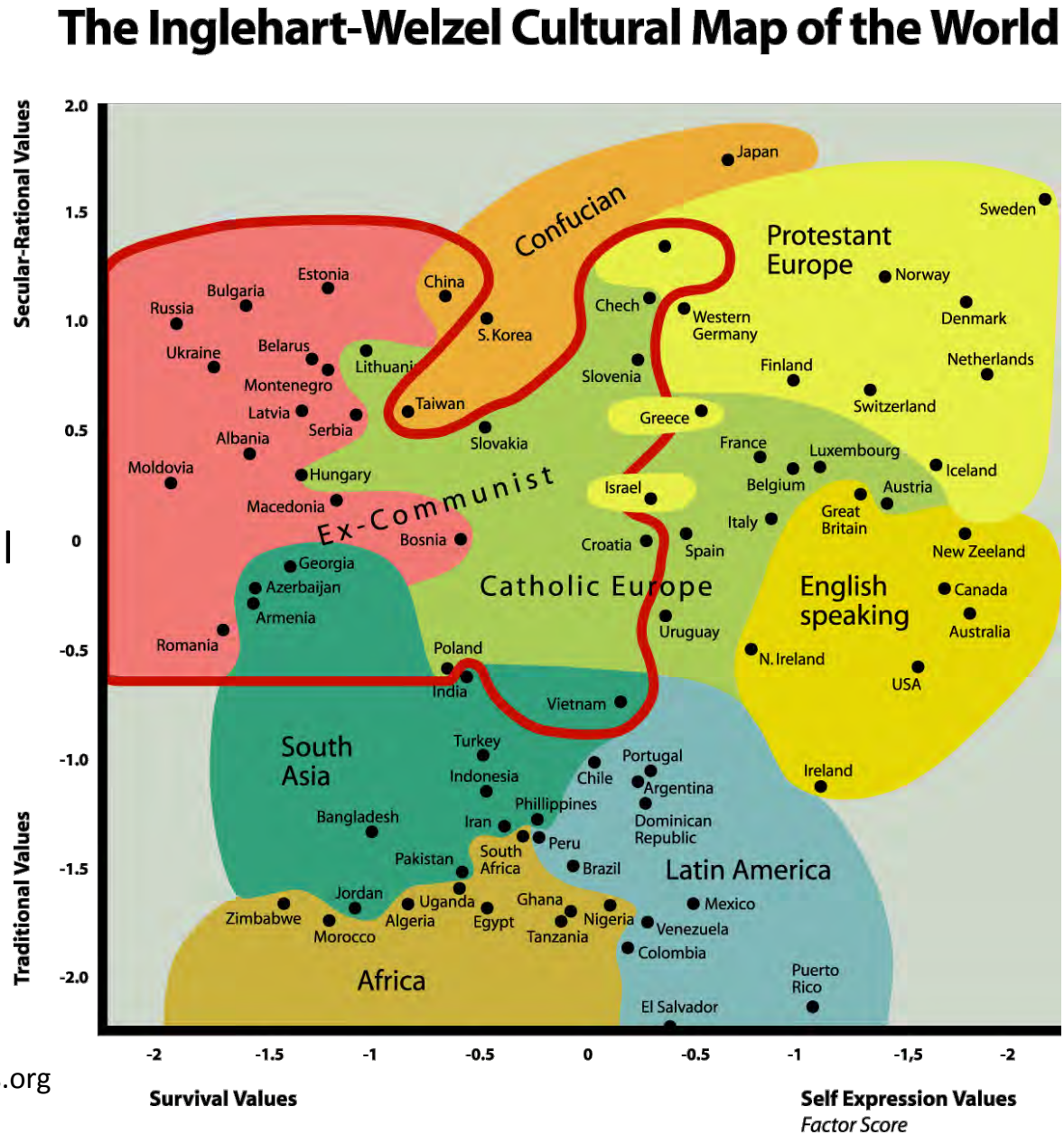
- 300-400 questions
- 30-70 countries
- +1.000 persons per country

Two scales explains 70% variation in Country Mean (Factor Analysis):

Traditional→Secular-rational values

&

Survival→Self-expression values



Source: Inglehart & Welzel 1997 + www.wvs.org

NORDIC METRICS

Parents values/influence → Important for the educational system

"Which qualities, if any, do you consider to be especially important that children are encouraged to learn at home"

	Mexico	Denmark	US	S. Korea
INDEPENDENCE	**39%**	**81%**	61%	77%
HARD WORK	38%	**2%**	61%	**72%**
FEELING OF RESPONSIBILITY	75%	81%	**72%**	**92%**
IMAGINATION	**24%**	**37%**	30%	33%
TOLERANCE & RESPECT FOR OTHERs	78%	**87%**	79%	**65%**
THRIFT SAVING MONEY AND THINGS	35%	**10%**	23%	**68%**
DETERMINATION & PERSEVERANCE	**27%**	32%	**45%**	42%
RELIGIOUS FAITH	35%	8%	**52%**	21%
UNSELFISHNESS	43%	**56%**	39%	**15%**
OBEDIENCE	**55%**	**14%**	32%	16%

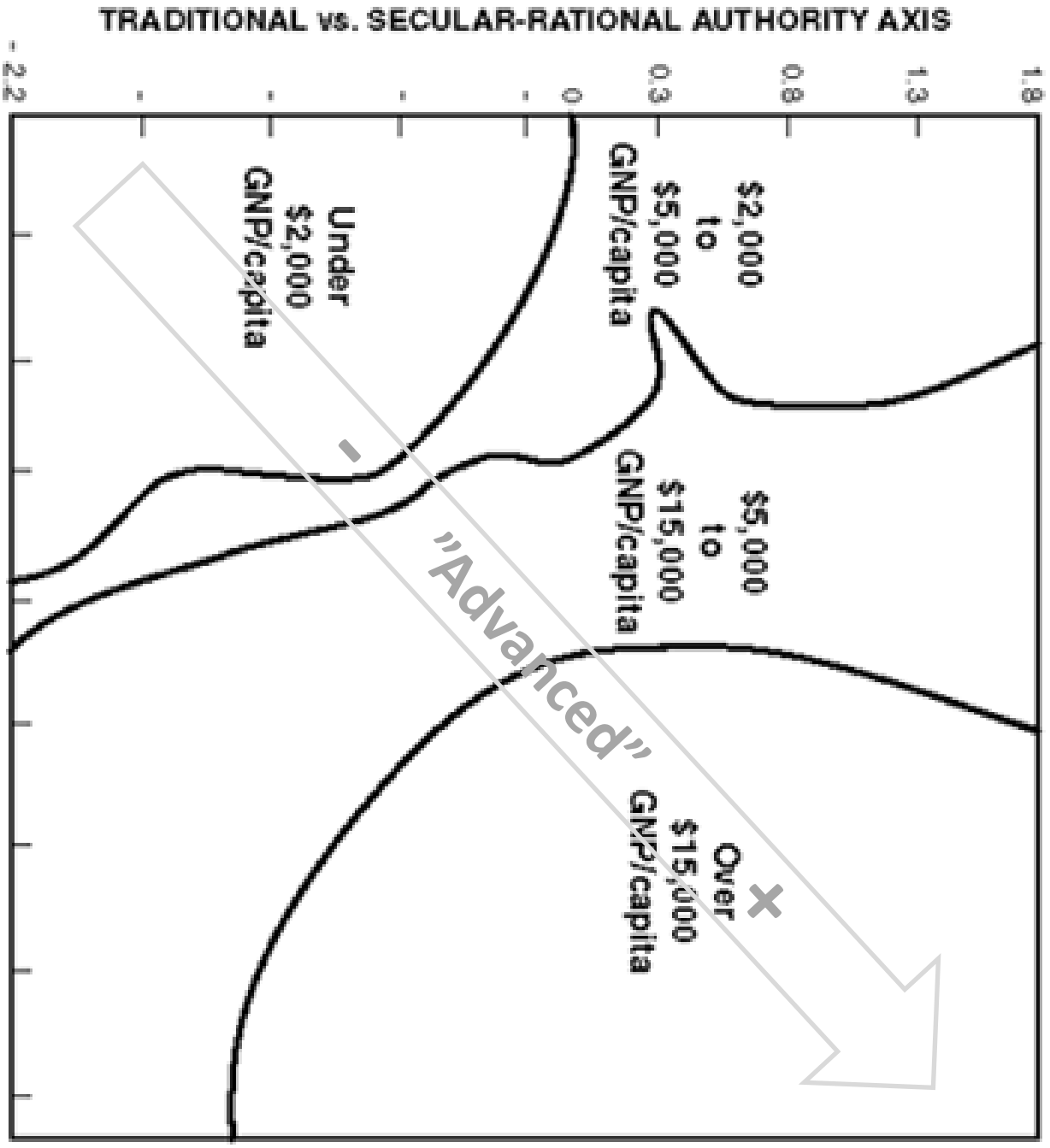
NORDIC METRICS

Different students' values

- Different challenges for the teachers

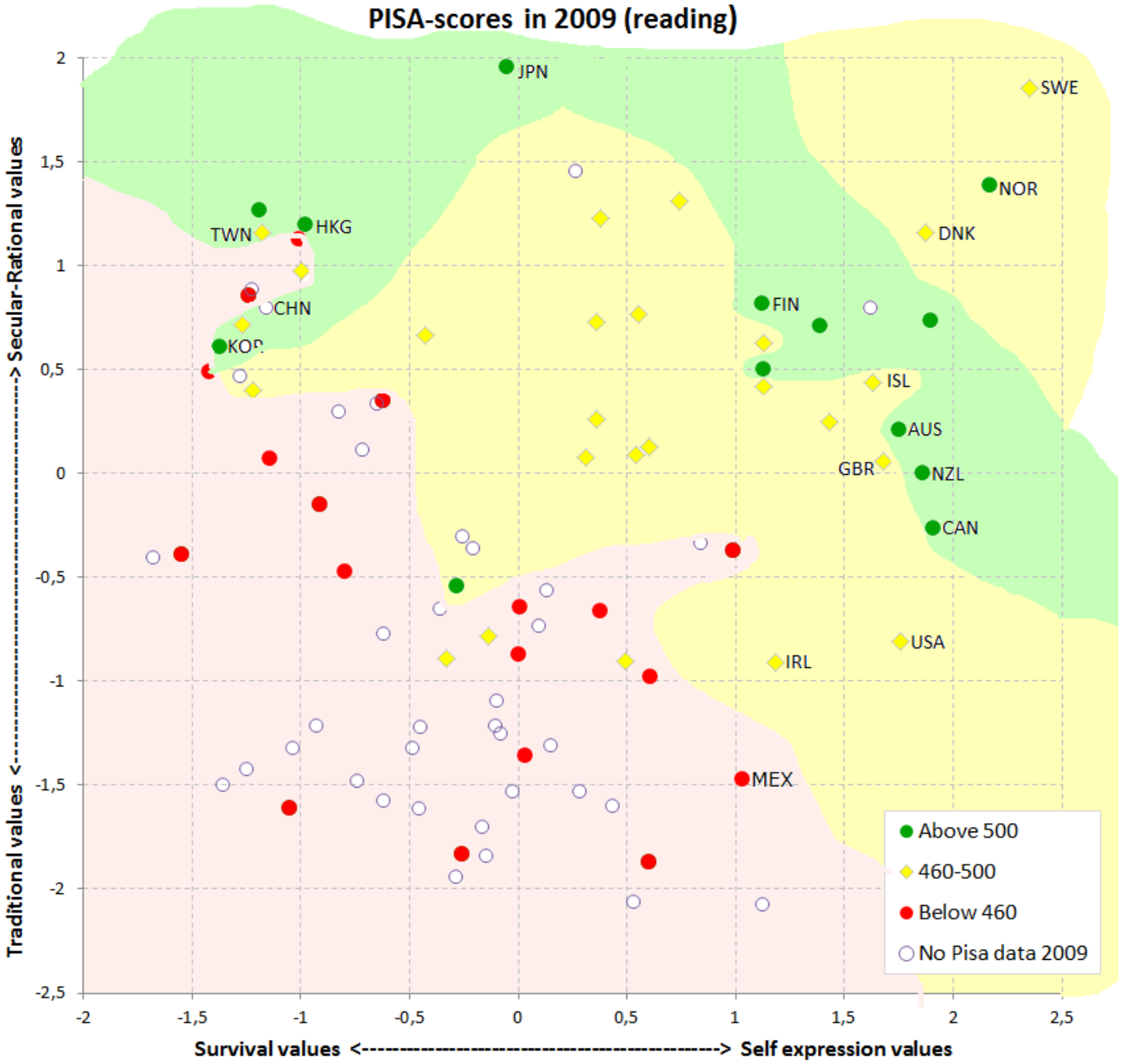
Secular-rational	**Confucian** + Hard work, Thrift, Responsibility - Tolerance, Unselfishness	**Nordic** + Independence, Imagination, Tolerance, Unselfishness - Hard work, Thrift, Obedience
Rational-Traditional		**Anglo-American** + Determination, Perseverance, Religious faith, - Responsibility
Traditional		**Latin-American** + Obedience - Independence, Determination, Imagination
	Survival	**Self-expression**

Economy & Values



NORDIC METRICS

Education (Reading/PISA) & Values



Formal Assessment in DK in the Basic school

NORDIC METRICS

2 generations of assessment in DK

- **1: Exams (1814→):**
- **2: National testing (2010→):**

NORDIC METRICS

Final exams (1814→)

- Assessment ***of*** learning
- Purpose: Summative / Information for the system
- Low-Medium Stakes
- Types of examinations:
  - Verbal, practical
  - Written/essays
  - Test-like
  - Mix of the above
- Items not tested / no fit to scales / not comparable over time / ***but they do test the full curriculum***
- NOW: Motivation and ceremoniel purposes
- FUTURE: Access to youth education (Stakes rising)

National testing system (2010→)

- Assessment ***for*** learning
- Purpose: Formative / Information for the teacher
- Low Stakes

NORDIC METRICS

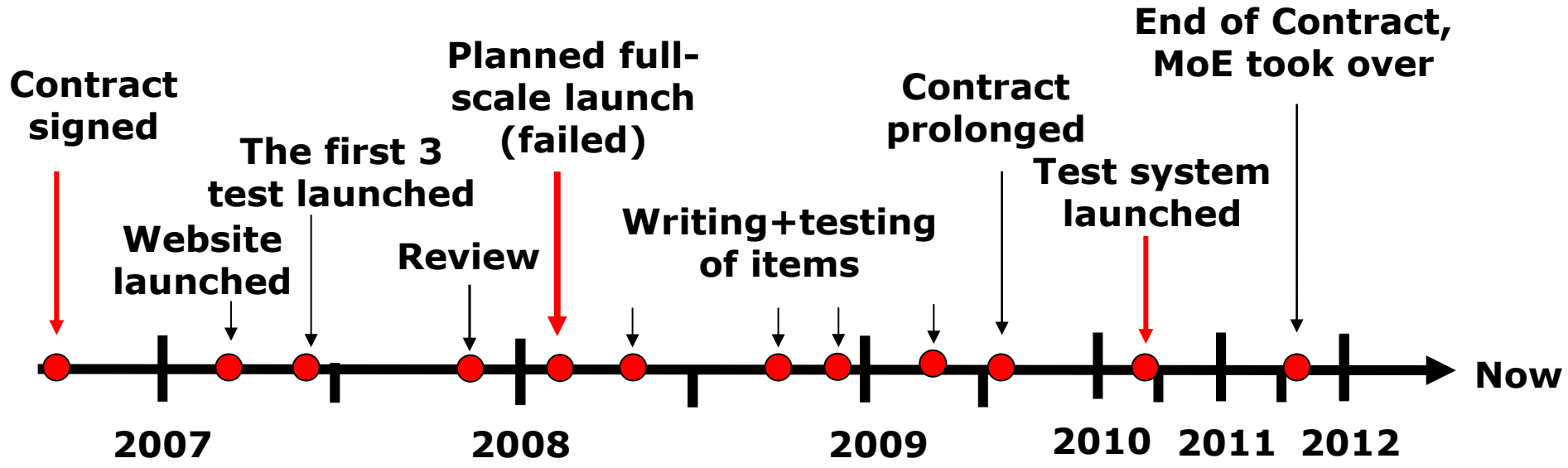
Background for the Danish initiatives to improve the culture of evaluation

1. PISA-surveys (2000-2003)
2. OECD-review on Denmark (2004) recommending mandatory exams and a National testing system
3. Governments initiatives according to the OECD-Recommendations

NORDIC METRICS

Development and implementation

- 2004-2006 Organization, Development, Tender & Contract.
- The first 3 test was launched 2007 (with reduced item banks)
- Sep. 2007, An expert review was conducted that showed, that
  - the psychometric standard of the system was at a high level, but
  - the quality of the items and size of the item pools were both insufficient
- 2008-2009, Development and testing of items in
- Autumn 2009, Full scale tests of the test system
- March 2010, Test system launched
- Aug. 2011, Contract ended, no more money/development, MoE agency took over
- Sep. 2013, Policy evaluation of the test system → Very limited formative use
- 2013→ : Local municipalities develop there own initiatives → Formative assessment



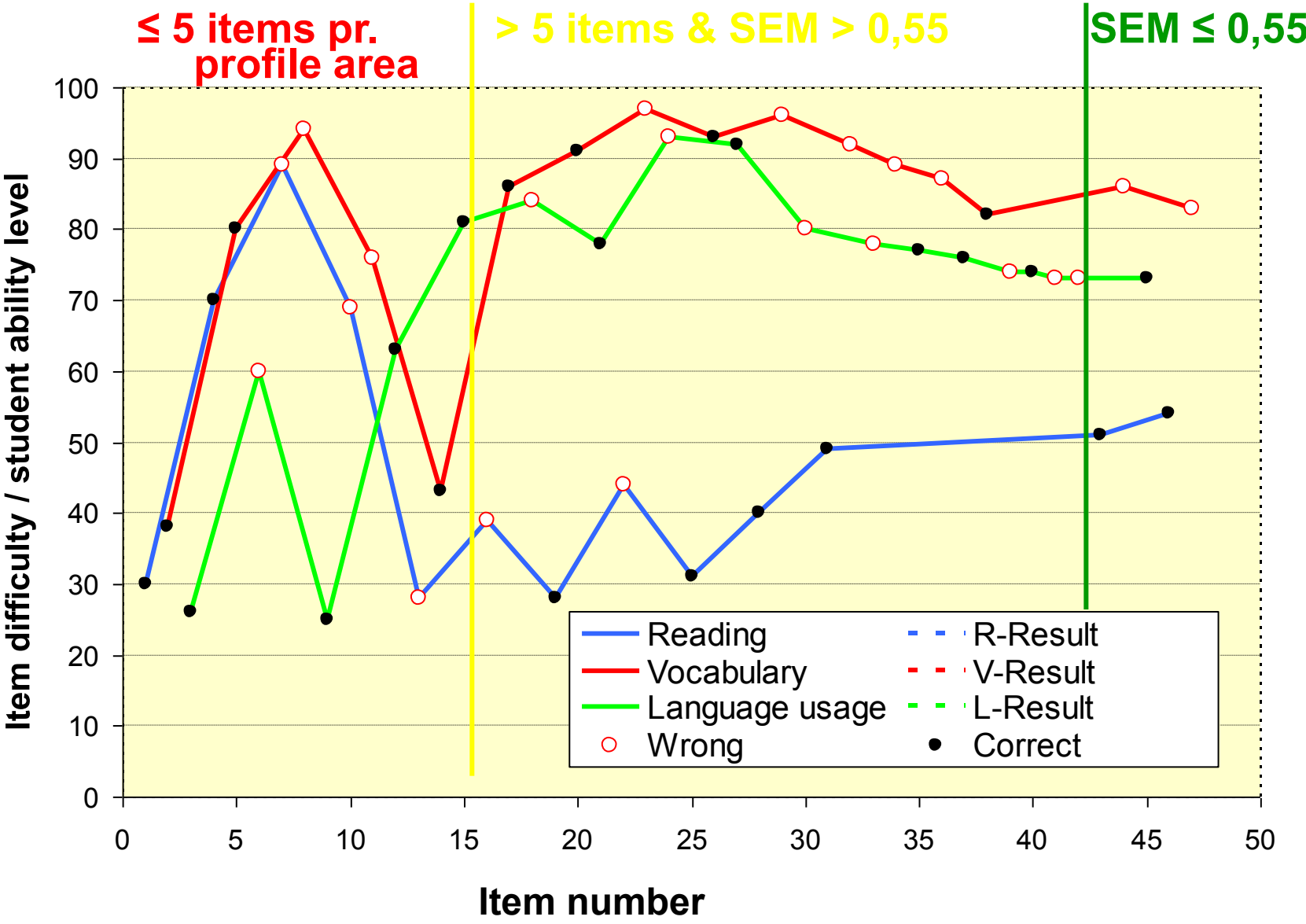
NORDIC METRICS

The National testing system (2010→):

- Full scale implemented in spring 2010
- Adaptive testing:
  - 10 mandatory test : Reading (2-4-6-8), Math (3-6), English (7) Physics, Geography & Biology (8)
  - 3 contend areas in each test,
  - Strictly formative (results are confidential)
  - Administering: By teacher - MoE provide guidance (mostly followed), but the teacher makes the decision. Not completely comparable but fairly reliable aggregated data.
- Assessment of students – not of teachers/schools!!
- First tool for measurement of reading on a scale in DK
- Lots of schools don't know what to do with the results

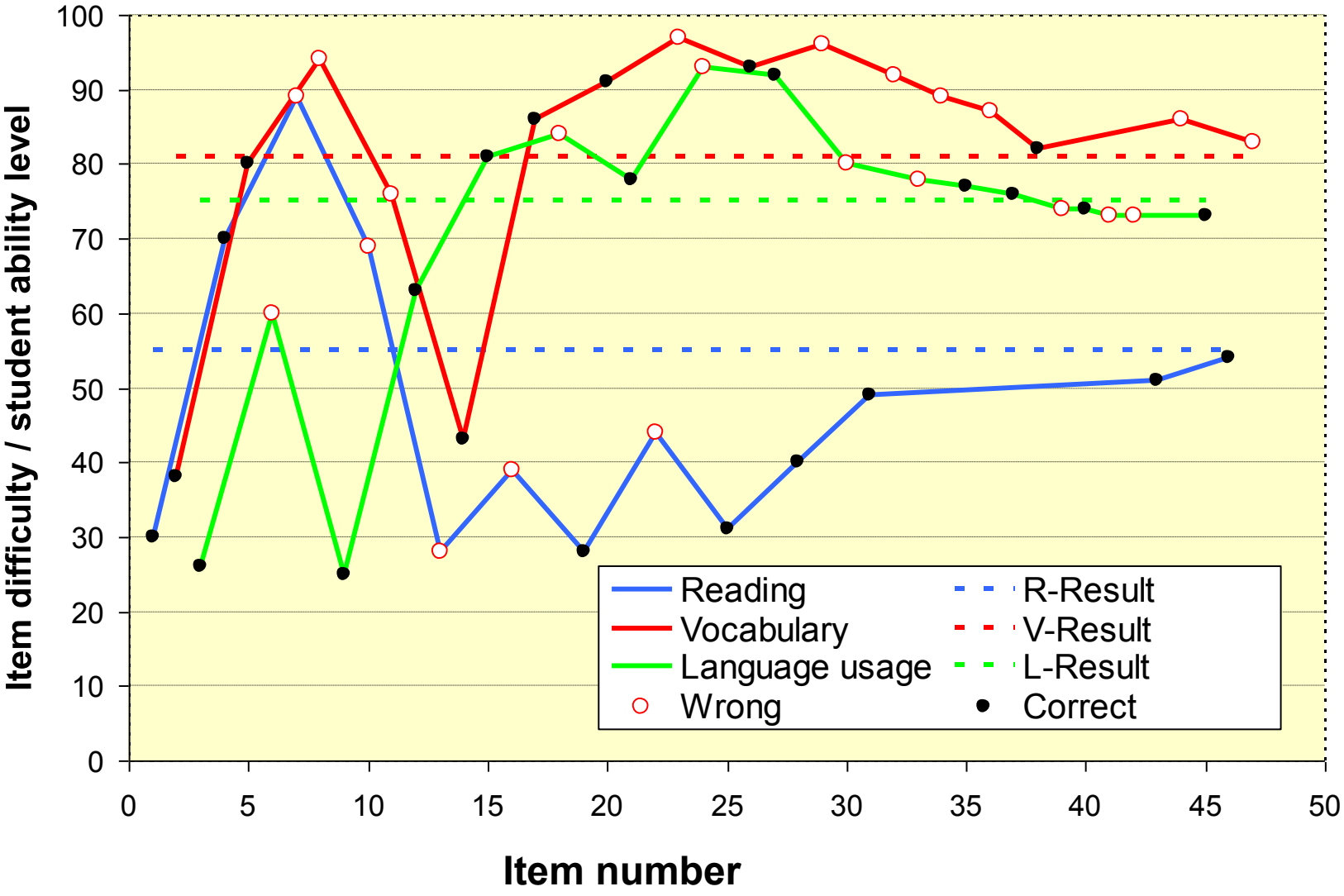
NORDIC METRICS

Adaptive test – 3 profile areas



NORDIC METRICS

Adaptive test – 3 profile areas



NORDIC METRICS

CAT (Computer Adaptive Testing)

- Adaptive = Adapts to the individual students ability:

  Correct answer ➔ More difficult questions
  Wrong answer ➔ Easier questions

  The test is most efficient when item difficulty = student ability

- More effective testing ➔ more detailed results:

  Adaptive within 3 “profile areas” ➔ 3 tests in 1

  (E.g. English: Reading, Vocabulary and Language usage)

- Simple principles, but a few tricky conditions:
  - Extensive demands to the technology – both capacity and stability.
  - Very large Item-pools with the right mix of high quality items

NORDIC METRICS

The national tests

Subject \ Form	1	2	3	4	5	6	7	8	9
Danish / reading		X		X		X		X	
Math			X			X			
English							X		
Geography								X	
Biology								X	
Physics/chemistry								X	
Danish as 2'nd language									

Forms where tests can be used

X 10 compulsory tests

Requirement for the item bank

Condition for the Danish national adaptive testing is:

- That the difficulty of the items are well defined and stabile over time (unidimensional & homogeneity).
- No differential item function.
- That the items are evenly distributed on item-difficulty → challenges for all students.
- Sufficient number of items so even the fastest student doesn't empty the items bank
- Students meet mostly new item, when tests are used more than once (max 3 test courses pr. student/test, including compulsory test)

Therefore following was required in the contract:

- Minimum 540 items pr. test (180 pr. Profile area) – evenly distributed on difficulty levels
- All items are tested on 500-700 students
- The items are required to fit a Rasch model

Different types of questions/items

- A variety of formats are used but no open ended items
- Both dichotomous and polytomous items are used (the Rasch partial credit model)
- Polytomous items/partial credit provides opportunities for further development. E.g. integrating the time dimension (still unidimensional) in the estimation of the students reading/decoding ability.

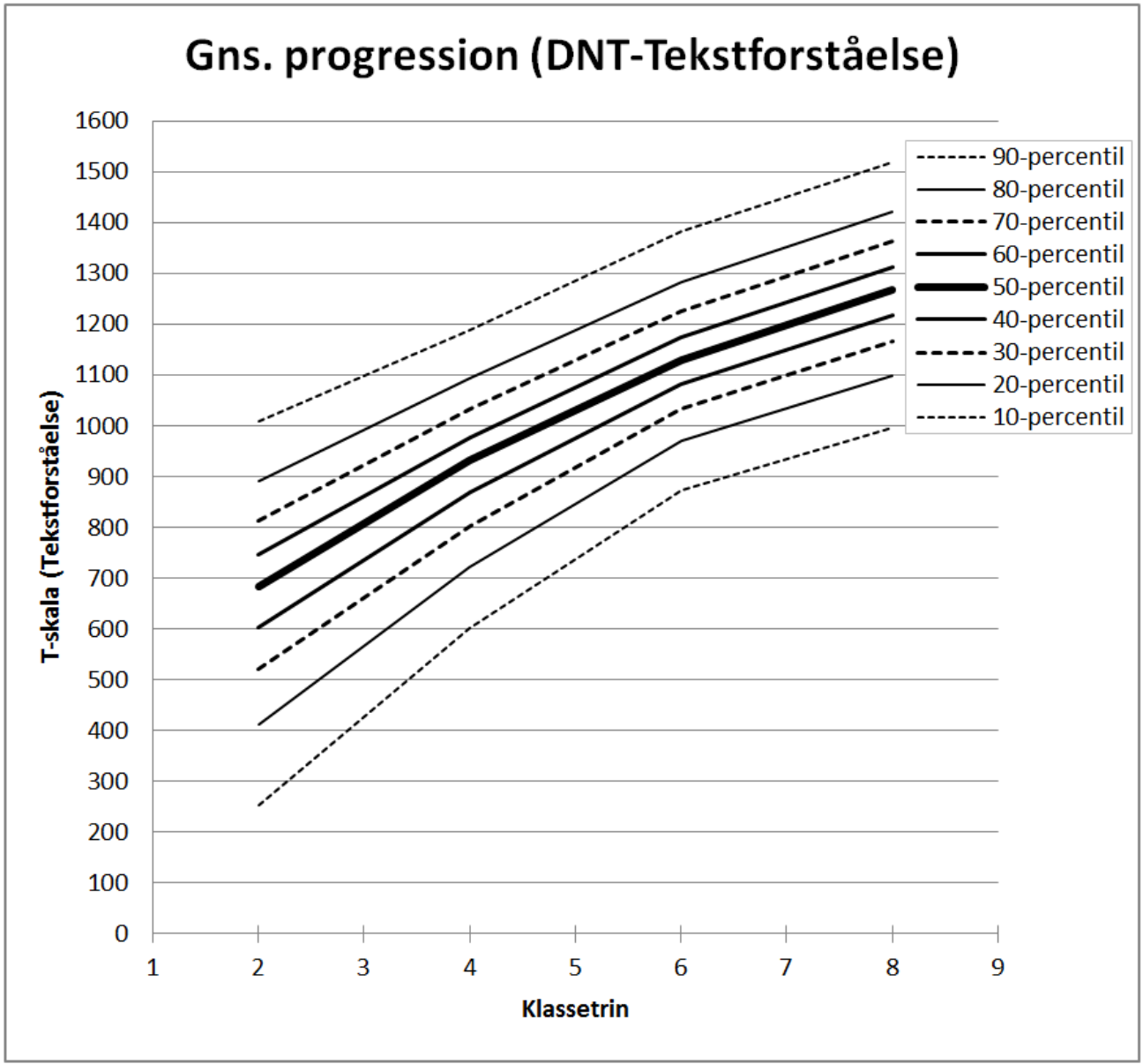
NORDIC METRICS

Advantages of adaptivity: Targetted testing

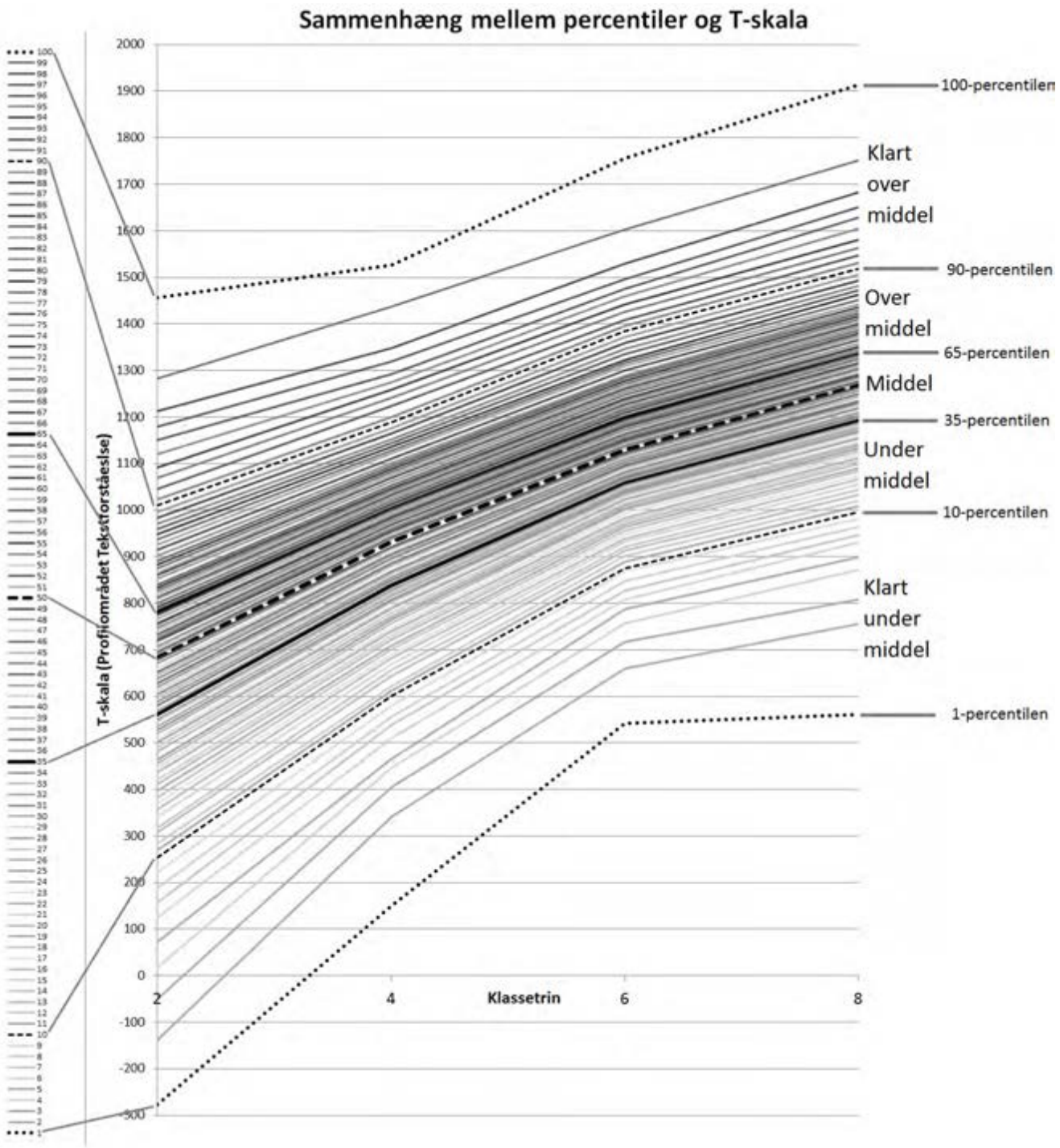
- The adaptive model performs relatively best in the extremes (measuring proficiency of high and low achievers).
- Reduction of SEM to 0,55 (reliability app. 0,7) normally within 8-15 items ***independently of student ability.***
- The adaptive test is generally a good experience for the low achievers, but can be a frustrating experience for high achievers (no free rides).
- The greater the differences in ability within the tested group of students, the greater the advantage of adaptive testing.
- A comprehensive school with significant academic differences in every classroom

Advantages of adaptivity: Targetted testing

Distance between 10% best and 10% worst performing students in the class: 6-8 year's education



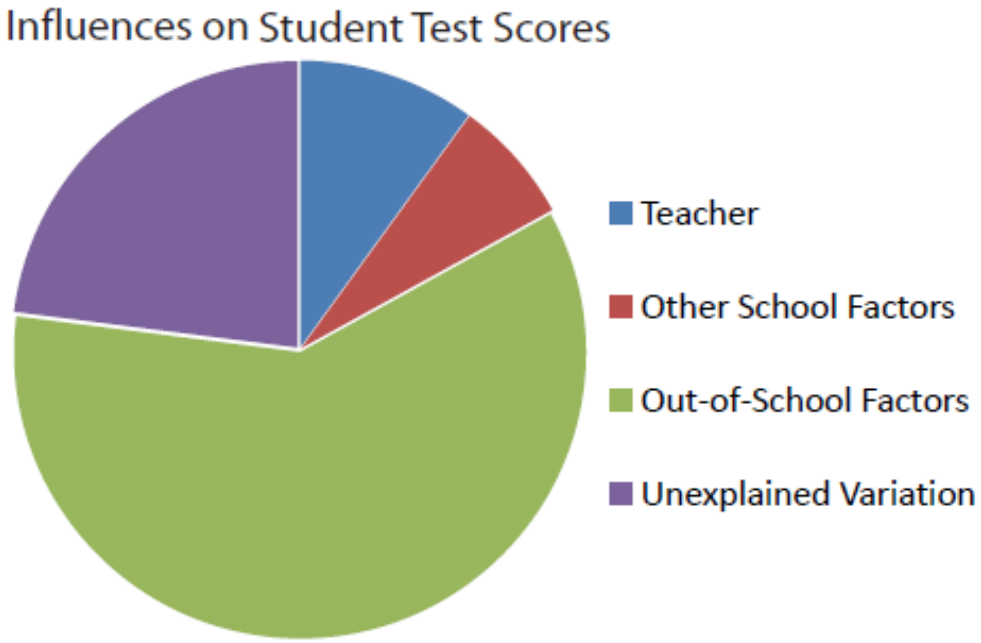
NORDIC METRICS



NORDIC METRICS

Schools have some effect on Growth, the student background determines the final outcome

. . . in US?



Haertel, E., 2013, (Retrived 26052014
https://www.ets.org/Media/Research/pdf/PICANG14.pdf)

In DK . . . . .

Final test scores		School performance (growth/added value)	
		Low	High
Student background (mean score – entrance)	Low	–	-
	High	+	+

NORDIC METRICS

Administering and getting results

- Different results on various levels (teacher, headmaster, municipality) on a need to know basis (avoid ranking)
- The teacher gets results online for
  - the class,
  - the students scores and
  - Details of the tests, including access to the items/respnses in the individual tests.

NORDIC METRICS

Screen 1: Class - Distribution of results (5 point scale based on percentiles)

Name: **Ole Jorgensen (olex2561)**

School : **Evaluating School**

Login Status: **Teacher**

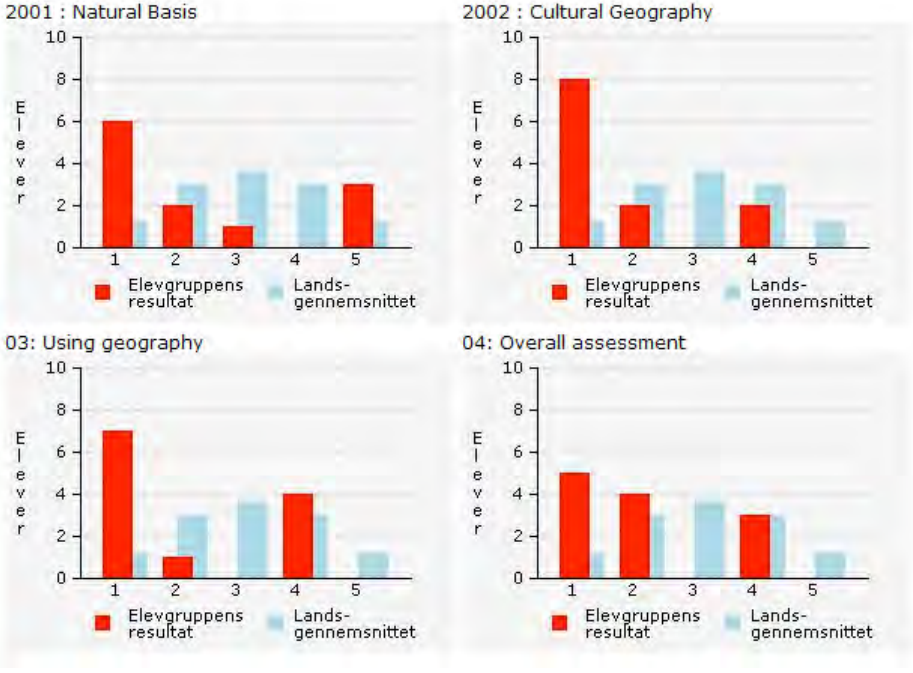
- **Booking**
- **Test Overview**
- **Status**
- **Test Results Management**
- **Test Results**
- **> > 3.U Mathematics 3rd at .**
- **> >8a Geography 8th at .**
- **> > All students - Tables**
- **> > Details pupils**
- **> > Print to home**
- **> > Create / student groups**
- **To the front page of the website**
- **Manuals**
- **SUPPORT**

Contacting UNI • C
tel 89 37 66 00
Mon .- two . at . 8-17
fr . at . 8-15
support@uni-c.dk

Confidential

Information on the test tasks and test results are confidential, see primary Law § 55 b. The reason - the failure to print and store non screen
Unauthorized disclosure or use of test results and test tasks will be punishable under the Penal Code provisions on professional secrecy .

Period: Last test results
Class: 8a
Test : Geography 8th at .
Teacher : Ole Jorgensen (olex2561)
Number of students who have completed : 12



NORDIC METRICS

Screen 2: Class – Student results (percentiles)

NATIONAL TEST

Name: **Ole Jorgensen (olex2561)**

School : **Evaluating School**

Login Status: **Teacher**

- Booking
- Test Overview
- Status
- Test Results Management
- Test Results
- > > 3.U Mathematics 3rd at .
- > >8a Geography 8th at .
- > >All pupils - Tables
- > > Details pupils
- > > Print to home
- > > Create / student groups
- To the front page of the website
- Manuals

SUPPORT

Contacting UNI • C
tel 89 37 66 00
Mon .- two . at . 8-17
fr . at . 8-15
support@uni-c.dk

Results and status - Table

Confidential

Information on the test tasks and test results are confidential, see primary Law § 55 b. The reason - the failure to print and store non screens. Unauthorized disclosure or use of test results and test tasks will be punishable under the Penal Code provisions on professional secrecy .

Period: Last test results
Class: 8a Test : Geography 8th at .
Teacher : Ole Jorgensen (olex2561)
Number of students : 12

Sort by pressing headlines

Name	Unilogin	Natural Basis Score 1-100	Cultural Geography Score 1-100	Using geography Score 1-100	Overall rating Score 1-100	Status
Bo Jensen	boxx1548	97 = Well above average	89 = Above Average	83 = Above Average	90 = Above Average	Completed
Ditlev Jensen	ditl0096	2 = Well Below Average	1 = Well Below Average	69 = Above Average	24 = Below Average	Completed
Ejnar Jensen	ejne0047	58 = Average	90 = Above Average	72 = Above Average	73 = Above Average	Completed yellow
Erik Eriksen	erik2118	1 = Well Below Average	1 = Well Below Average	1 = Well Below Average	1 = Well Below Average	Completed
Frederik Jensen	fred289c	13 = Below Average	6 = Well Below Average	1 = Well Below Average	7 = Well Below Average	Completed
Hans Jensen	hans2914	30 = Below Average	7 = Well Below Average	1 = Well Below Average	13 = Below Average	Completed yellow
Ida Jensen	idax7005	98 = Well above average	12 = Below Average	87 = Above Average	66 = Above Average	Completed
Jane Jensen	jane2338	4 = Well Below Average	19 = Below Average	32 = Below Average	18 = Below Average	Completed
Mia Jensen	miax5706	1 = Well Below Average	1 = Well Below Average	1 = Well Below Average	1 = Well Below Average	Completed
Oline Jensen	olin0149	94 = Well above average	1 = Well Below Average	1 = Well Below Average	32 = Below Average	Completed yellow
Svend Petersen	sven1138	7 = Well Below Average	1 = Well Below Average	2 = Well Below Average	3 = Well Below Average	Completed
Svend Jensen	sven1232	5 = Well Below Average	2 = Well Below Average	1 = Well Below Average	3 = Well Below Average	Completed
Class Average		34 = Below Average	19 = Below Average	29 = Below Average	28 = Below Average	

Screen 3: Individual student test course (access to items & reponses)

NORDIC METRICS

NATIONAL TEST

Name: **Ole Jorgensen (olex2561)**

School : **Evaluating School**

Login Status: **Teacher**

- Booking
- Test Overview
- Status
- Test Results Management
- Test Results
- > > 3.U Mathematics 3rd at .
- > >8a Geography 8th at .
- > > All students - Tables
- > >Details pupils
- > > Print to home
- > > Create / student groups
- To the front page of the website
- Manuals

SUPPORT

Contacting UNI • C
tel 89 37 66 00
Mon .- two . at . 8-17
fr . at . 8-15
support@uni-c.dk

Answer

Confidential

Information on the test tasks and test results are confidential, see primary Law § 55 b. The reason - the failure to print and store non screens. Unauthorized disclosure or use of test results and test tasks will be punishable under the Penal Code provisions on professional secrecy .

Questions	Profile Area	Task Topic	Score	Difficulty	Nominal duration	Spent time	TestAssignments
1	Language Understanding	Verbs	F	***	12 sec.	25 sec.	View job
2	Decoding	Compound words	R	*	15 sec.	24 sec.	View job
3	Text Understanding	Could form inferences in a literary text (Inferensdan	R	**	51 sec.	63 sec.	View job
4	Language Understanding	Fixed terms	R	**	14 sec.	16 sec.	View job
5	Decoding	Bent / derivative lydrette words	R	**	23 sec.	34 sec.	View job
6	Text Understanding	Identify the text's topic / idea in a sagprosatekst	R	***	38 sec.	65 sec.	View job
7	Language Understanding	Fixed terms	R	***	13 sec.	13 sec.	View job
8	Decoding	Compound words	R	**	25 sec.	19 sec.	View job
9	Text Understanding	Could find and summarize several factual informatio	F	****	34 sec.	41 sec.	View job
10	Language Understanding	Verbs	F	****	12 sec.	12 sec.	View job
11	Decoding	Compound words	R	***	20 sec.	37 sec.	View job
12	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	34 sec.	37 sec.	View job
13	Language Understanding	Adjectives	R	***	12 sec.	14 sec.	View job
14	Decoding	Words with conditional says	R	****	19 sec.	16 sec.	View job
15	Text Understanding	Could find specific information in a literary text	R	**	35 sec.	58 sec.	View job
16	Language Understanding	Fixed terms	R	****	13 sec.	15 sec.	View job
17	Decoding	Words with conditional says	F	******	28 sec.	40 sec.	View job
18	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	41 sec.	48 sec.	View job
19	Language Understanding	Nouns	F	****	17 sec.	23 sec.	View job
20	Decoding	Compound words	F	****	17 sec.	13 sec.	View job
21	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	R, R, R	***	64 sec.	131 sec.	View job
22	Language Understanding	Nouns	R	****	25 sec.	44 sec.	View job
23	Decoding	Compound words	F	***	32 sec.	39 sec.	View job
24	Text Understanding	Could form the inner relevant mental images (intern	F	***	55 sec.	34 sec.	View job
25	Language Understanding	Nouns	R	****	15 sec.	8 sec.	View job
26	Decoding	Compound words	R	***	24 sec.	17 sec.	View job
27	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	64 sec.	57 sec.	View job
28	Language Understanding	Nouns	R	******	13 sec.	12 sec.	View job
29	Decoding	Words with conditional says	R	***	16 sec.	12 sec.	View job
30	Text Understanding	Identify the text's topic / idea in a sagprosatekst	R	***	37 sec.	49 sec.	View job
31	Language Understanding	fixed expressions	F	******	13 sec.	15 sec.	View job

NORDIC METRICS

Problem: The teachers don't use the results formatively.

Technical solutions

1. Meaningful presentation of the results:
   - Prior: Snapshots, Now: Growth (added value)
   - Prior: Percentile scales, Now: Interval scales
   - Prior: Tables, Now: Graphics

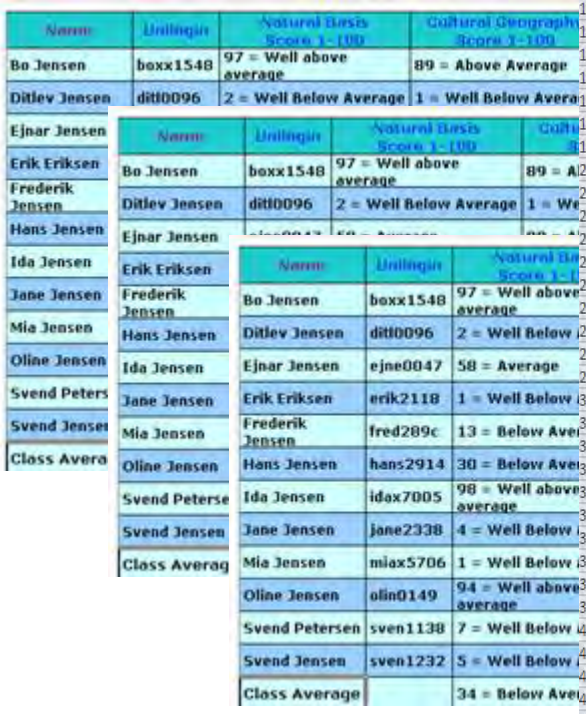
Teacher education

1. Teachers must be instructed in how to read and interpret test results – in a critical way to identify problems "response validity"
2. We must find out what works. We need evidence on effects of pedagogical methods.
3. Teachers must be able to combine knowledge from test results with inferences from pedagogical research to improve their instructional practice.

Test scores for same class in 2, 4 & 6 grade, Reading

Needed: Tool for flexible presentation of individual student growth results on interval scales

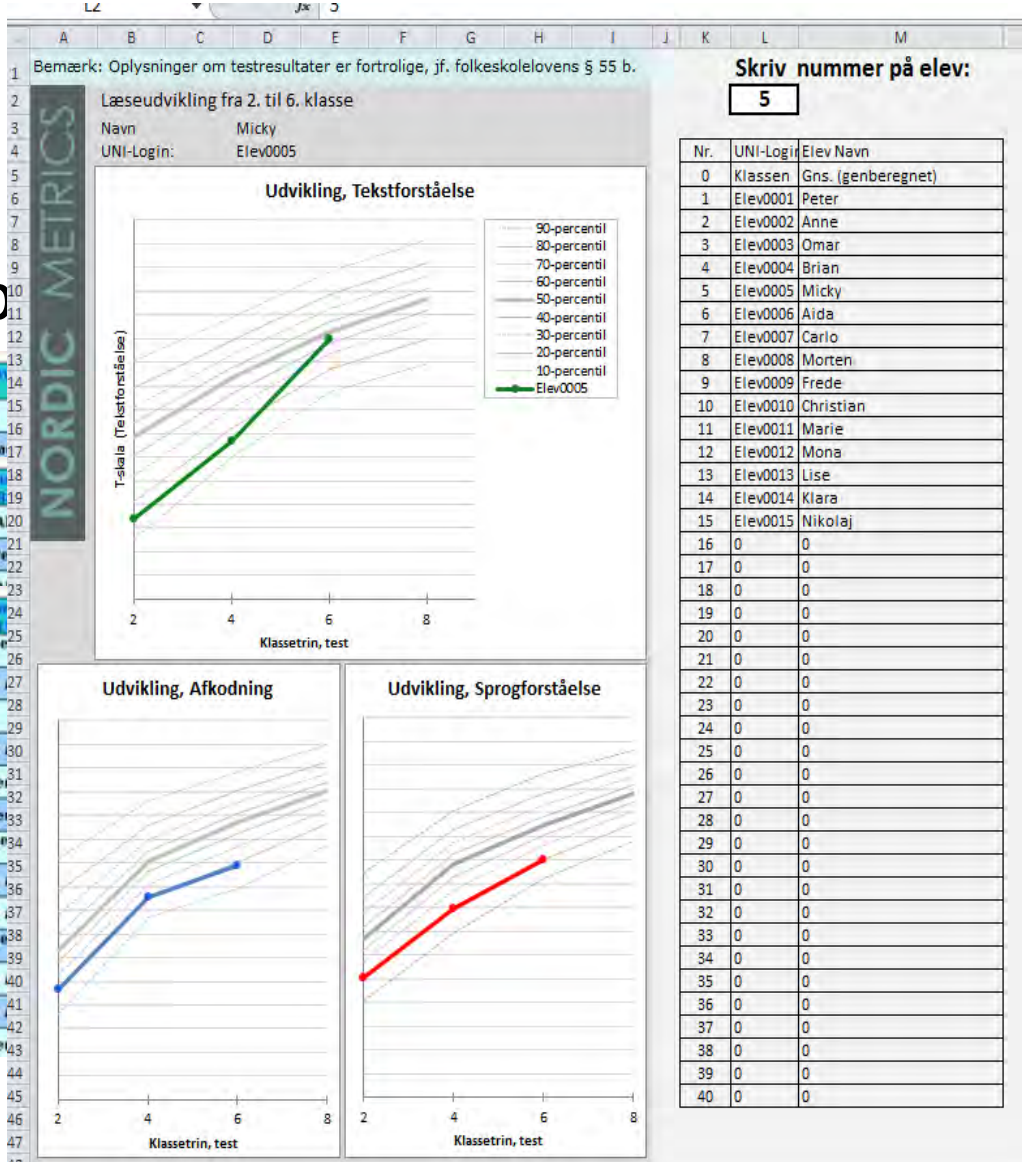
Available today: 3 p



Navn	Unilogin	Natural Basis Score 1-100	Cultural Geography Score 1-100
Bo Jensen	boxx1548	97 = Well above average	89 = Above Average
Ditlev Jensen	ditl0096	2 = Well Below Average	1 = Well Below Avera
Ejnar Jensen			
Erik Eriksen			
Frederik Jensen			
Hans Jensen			
Ida Jensen			
Jane Jensen			
Mia Jensen			
Oline Jensen			
Svend Peters			
Svend Jense			
Class Avera			

Navn	Unilogin	Natural Basis Score 1-100	Cultural Geography Score 1-100
Bo Jensen	boxx1548	97 = Well above average	89 = A
Ditlev Jensen	ditl0096	2 = Well Below Average	1 = We
Ejnar Jensen			
Erik Eriksen			
Frederik Jensen			
Hans Jensen			
Ida Jensen			
Jane Jensen			
Mia Jensen			
Oline Jensen			
Svend Peterse			
Svend Jensen			
Class Averag			

Navn	Unilogin	Natural Basis Score 1-100
Bo Jensen	boxx1548	97 = Well above average
Ditlev Jensen	ditl0096	2 = Well Below
Ejnar Jensen	ejne0047	58 = Average
Erik Eriksen	erik2118	1 = Well Below
Frederik Jensen	fred289c	13 = Below Ave
Hans Jensen	hans2914	30 = Below Ave
Ida Jensen	idax7005	98 = Well above average
Jane Jensen	jane2338	4 = Well Below
Mia Jensen	miax5706	1 = Well Below
Oline Jensen	olin0149	94 = Well above average
Svend Petersen	sven1138	7 = Well Below
Svend Jensen	sven1232	5 = Well Below
Class Average		34 = Below Ave



Bemærk: Oplysninger om testresultater er fortrolige, jf. folkeskolelovens § 55 b.

Læseudvikling fra 2. til 6. klasse

Navn: Micky

UNI-Login: Elev0005

Skriv nummer på elev: 5

Nr.	UNI-Login	Elev Navn
0	Klassen	Gns. (genberegnet)
1	Elev0001	Peter
2	Elev0002	Anne
3	Elev0003	Omar
4	Elev0004	Brian
5	Elev0005	Micky
6	Elev0006	Aida
7	Elev0007	Carlo
8	Elev0008	Morten
9	Elev0009	Frede
10	Elev0010	Christian
11	Elev0011	Marie
12	Elev0012	Mona
13	Elev0013	Lise
14	Elev0014	Klara
15	Elev0015	Nikolaj
16	0	0
17	0	0
18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	0	0
39	0	0
40	0	0

NORDIC METRICS

Student Test: The adaptive course in three profile areas & reading speed (Realized/Expected time)

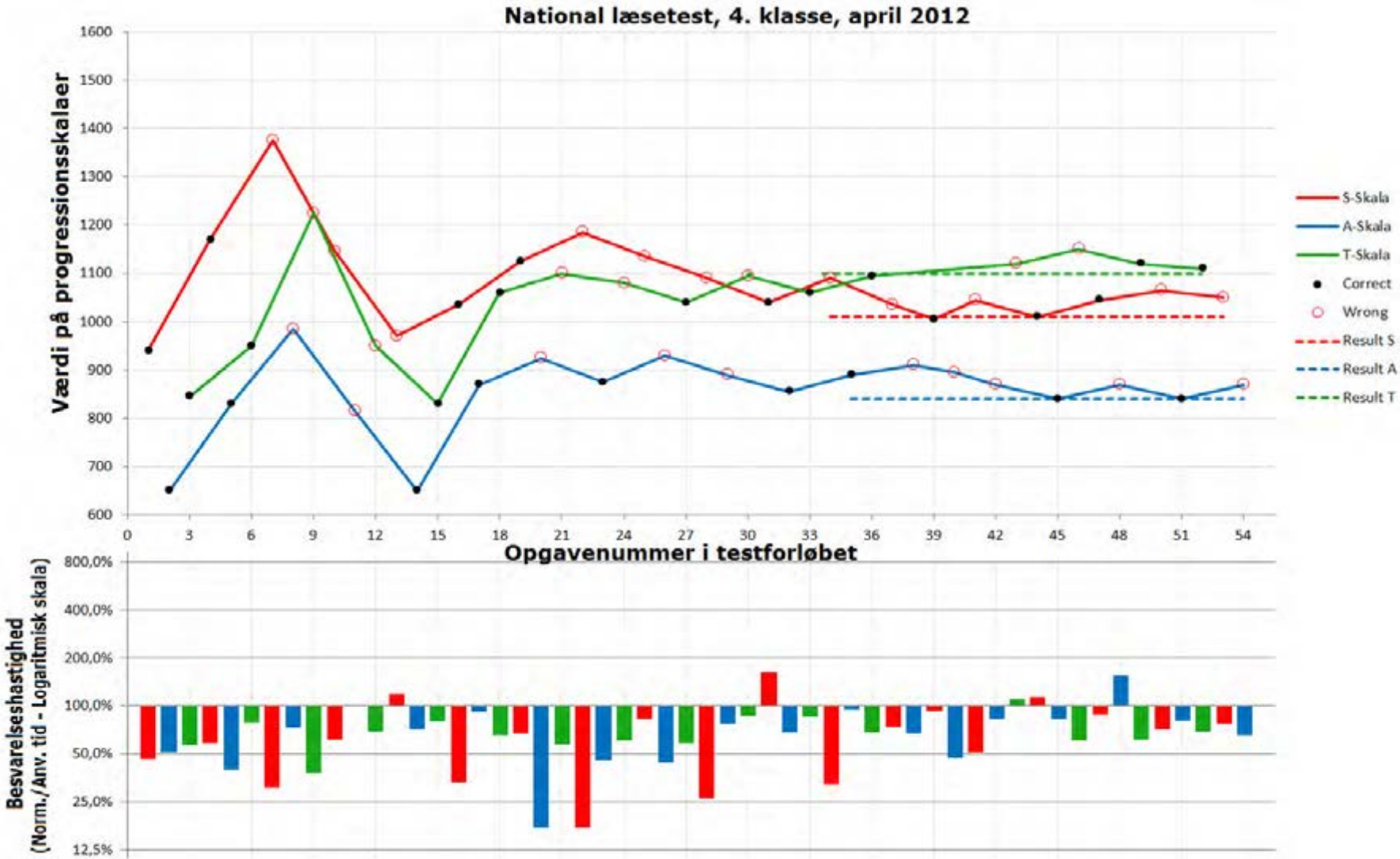
Available today: 6 dimensional table, 45-90 rows

Questions	Profile Area	Task Topic	Score	Difficulty	Nominal duration	Spent time	TestAssignments
1	Language Understanding	Verbs	F	***	12 sec.	25 sec.	View job
2	Decoding	Compound words	R	*	15 sec.	24 sec.	View job
3	Text Understanding	Could form inferences in a literary text (Inferensdani	R	**	51 sec.	63 sec.	View job
4	Language Understanding	Fixed terms	R	**	14 sec.	16 sec.	View job
5	Decoding	Bent / derivative lydrette words	R	**	23 sec.	34 sec.	View job
6	Text Understanding	Identify the text's topic / idea in a sagprosatekst	R	***	38 sec.	65 sec.	View job
7	Language Understanding	Fixed terms	R	***	13 sec.	13 sec.	View job
8	Decoding	Compound words	R	**	25 sec.	19 sec.	View job
9	Text Understanding	Could find and summarize several factual informatio	F	****	34 sec.	41 sec.	View job
10	Language Understanding	Verbs	F	****	12 sec.	12 sec.	View job
11	Decoding	Compound words	R	***	20 sec.	37 sec.	View job
12	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	34 sec.	37 sec.	View job
13	Language Understanding	Adjectives	R	***	12 sec.	14 sec.	View job
14	Decoding	Words with conditional says	R	****	19 sec.	16 sec.	View job
15	Text Understanding	Could find specific information in a literary text	R	**	35 sec.	58 sec.	View job
16	Language Understanding	Fixed terms	R	****	13 sec.	15 sec.	View job
17	Decoding	Words with conditional says	F	*****	28 sec.	40 sec.	View job
18	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	41 sec.	48 sec.	View job
19	Language Understanding	Nouns	F	****	17 sec.	23 sec.	View job
20	Decoding	Compound words	F	****	17 sec.	13 sec.	View job
21	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	R, R, R	***	64 sec.	131 sec.	View job
22	Language Understanding	Nouns	R	****	25 sec.	44 sec.	View job
23	Decoding	Compound words	F	***	32 sec.	39 sec.	View job
24	Text Understanding	Could form the inner relevant mental images (intern	F	***	55 sec.	34 sec.	View job
25	Language Understanding	Nouns	R	****	15 sec.	8 sec.	View job
26	Decoding	Compound words	R	***	24 sec.	17 sec.	View job
27	Text Understanding	Could form inferences in a sagprosatekst (Inferensd	F	***	64 sec.	57 sec.	View job
28	Language Understanding	Nouns	R	*****	13 sec.	12 sec.	View job
29	Decoding	Words with conditional says	R	***	16 sec.	12 sec.	View job
30	Text Understanding	Identify the text's topic / idea in a sagprosatekst	R	***	37 sec.	49 sec.	View job
31	Language Understanding	fixed expressions	F	*****	13 sec.	15 sec.	View job
32	Decoding	Bent / derivative lydrette words	R	***	26 sec.	17 sec.	View job
33	Text Understanding	Could form inferences in a literary text (Inferensdani	B	***	43 sec.	3 sec.	View job
34	Language Understanding	Adjectives	F	****	11 sec.	8 sec.	View job
35	Decoding	Loanwords	R	***	13 sec.	24 sec.	View job
36	Text Understanding	Could find and summarize several factual informatio	R	***	56 sec.	58 sec.	View job
37	Language Understanding	verb	R	****	14 sec.	10 sec.	View job
38	Decoding	Words with conditional says	R	***	14 sec.	13 sec.	View job
39	Language Understanding	Nouns	R	*****	11 sec.	12 sec.	View job
40	Decoding	Bent / derivative lydrette words	F	***	12 sec.	10 sec.	View job
41	Language Understanding	Adjectives	F	*****	17 sec.	10 sec.	View job
42	Decoding	Bent / derivative lydrette words	R	***	14 sec.	9 sec.	View job
43	Decoding	Bent / derivative lydrette words	F	***	19 sec.	15 sec.	View job
44	Decoding	Compound words	R	***	25 sec.	12 sec.	View job

NORDIC METRICS

Student Test: The adaptive course in three profile areas & reading speed (Realized/Expected time)

Needed: Simple overview/illustration of what happend during the test



Lessons Learned

A. Separation of summative and formative assessment

1. Final examinations for accountability, summative, state administered, rise in stakes
   - No advanced psychometric modeling, state sets the rules (time, aids, proctoring).
   - Full curriculum tested, mixed methods, comparability limited.

2. National CAT as a pedagogical tool , formative, low stakes
   - Non-curricular general constructs – like PISA (10% of DK-curriculum tested), IRT modelling, unidimensional scales,
   - Administering: MoE provide guidance (mostly followed), but the teacher makes the decision. Not completely comparable but fairly reliable aggregated data.
   - Results confidential by law (no public ranking, no cheating).

Lessons Learned

B. Test can improve student learning only if results are used formatively, understood and used constructively by teachers:

1. Teachers must be convinced that improvement is possible - Good data can prove that students are able to learn more (**V**)
2. Results must be presented in formats that can be explained to and understood by teachers (**WIP**)
3. Teachers must be instructed in how to read and interpret test results – and identify problems (**WIP**)
4. We must find out what works. We need evidence on effects of pedagogical methods.
5. Teachers must be able to combine knowledge from test results with inferences from pedagogical research to improve their instructional practice.

Lessons Learned

C. Realistic expectations and objectives have to be established in dialogue with schools

1. Success measured by test scores is done by aggregating: It all comes down to the individual students progress.
2. Teachers generally do their best. Change rarely come about without information and guidelines to teachers of what to do differently.
3. If the teachers don't understand what to do & commits to the plan, the plan is deemed to fail.
4. Any success can look like a failure if the bar is set to high. And stakeholder alignment is difficult.
5. In implementation of educational policy is patience a necessary virtue

Lessons learned

D. Look outside school: Assessment embedded in training & other learning platforms than school

- New it-based tools (assessment ***as*** learning) are being developed. Purpose: Monitoring learning, adapt to student ability, real-time feedback → Enhance learning & motivation

  2 examples:

  - **Edsphere** (American, Lexile framework) – Reading training WEB-app. Draws on almost infinite amounts of authentic text. Is able to auto generated items fitting the rasch model. Measures / Adapts to students reading level. Reports results to both student and teacher. Only in English but a Spanish edition is planned.
  - **LABSTER** (Danish) A 3-d Simulation of Biotech Laboratory for Education in science. Embedded questions and feedback. Linear versions are on marked now and a project integrating adaptivity/measurement is being started up (supported by the Danish state). Only in English.
- Much is determined in age 0-6. This should be looked into.