

Improving the Quality of Online Tests and Assessments

Svend Andreas Horgen

Greta Hjertø

Jarle Larsen

Sør-Trøndelag University College (HiST)

Trondheim, Norway

The binary sudoku

EASY

0	1
	0

A LITTLE BIT
CHALLENGING

1	0

REALLY
DIFFICULT

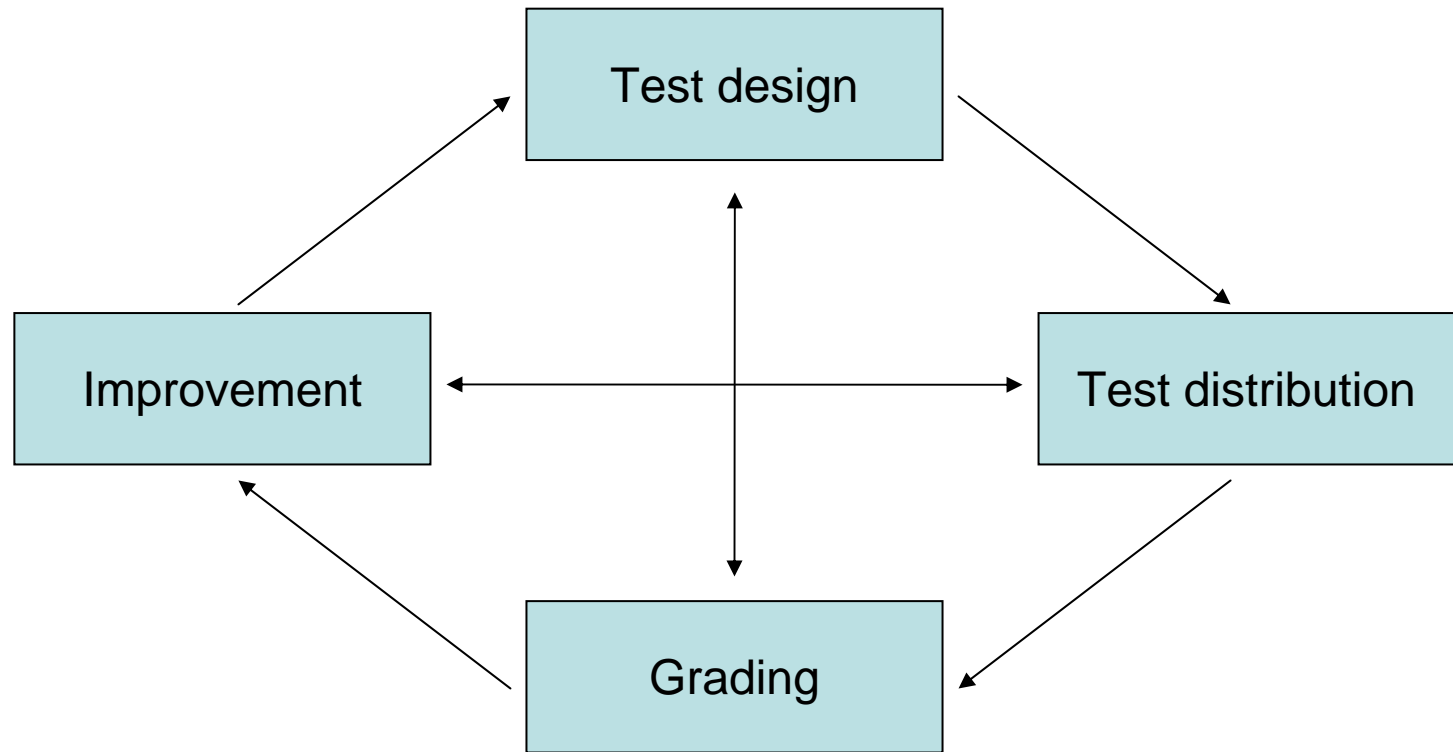
	1

PRETTY IMPOSSIBLE

1	1

Motivation

- A **quality reform** in higher education in Norway
- Distance education: 1000++
- Efficiency
- LMS has tools for testing, but
 - proprietary
 - lacks functionality
 - difficult to influence development



- 4 iterative phases for making good tests
 - Design: how to write high-quality questions, feedback strategy...
 - Distribution: delivery, cheat, # of attempts, ...
 - Grading: calculation of score, grading strategy, ...
 - *Improvement*: analysis, sort out bad questions, ...

Analysis, quality, success

- How should the scores be set?
 $x \rightarrow (x-100/n)*n/(n-1)$ Normalising, n is the number of alternatives, x is score
- Analysis of test results is important, but difficult and time consuming:
 - sorting, comparing
 - "popular" alternatives
 - question statistics
 - completion time
 - min, max, mean, frequency
 - distractor analysis
 - variances, standard deviation
 - other "statemactical" formulas

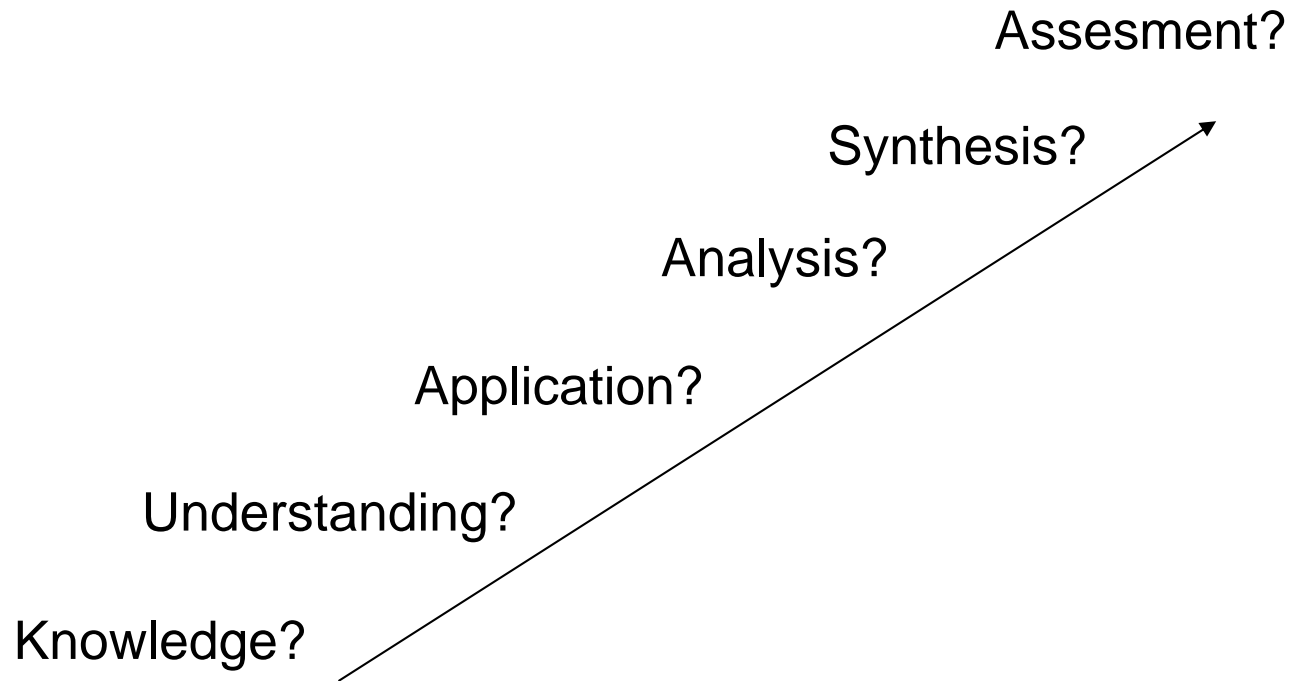
Implementation of a test tool

- Computer can automate and assist
 - easy development of tests and questions
 - reuse of questions and tests
 - administration of students and test distribution
 - random pick of questions
 - calculation of scores
 - analysis of results → statistics and proposals
- Computer tests are highly scalable

Web-based system: EVATEST

- All four phases
 - Focus on the improvement phase
- Multiple choice \leftrightarrow free text
- LaTeX code \rightarrow images of formulas
- Resources
 - files (images etc)
 - links (www)
- Test parameters
 - grading strategy
 - timing
 - availability
 - # of attempts
 - response strategy
 - ...
- IMS QTI: import and export
- Question pool

Blooms taxonomy – what can be tested?



Question pool

- All questions are saved in a common pool
- Searchable
 - keyword
 - free text
 - category
 - question type
 - author
 - subject
- Reuse and statistics...
- The tool knows the history of each question
- Resource for the teacher

Search result in question pool

Evatesting

http://elite.stud.aitel.hist.no/evatesting/index.jsp?meny=test&testvalg=visTester&handling=rediger&oppgave=leggitilloggavefrapool&testid=-10&leggitil=&lagre=lagre

Side-søk

Søk etter neste Tale Spol tilbake Tilbake Framover Hurtig framover Oppdater Åpne Lagre Skriv ut

Innlogget som: Svend Andreas Lærer:

Main page Subject Test Testcopying Solutions Statistics Taskpool Personalinfo Chat Help Logg out

Here you can add tasks direct in to the pool

Texttask Yes/no task Multiple-choice Find pair QTI/IMS import Edit pool

Help (Open in a new window)

Here you can search for tasks in the pool, look at them or delete them

Type: All types of tasks Task: Headword: mengd Author: truls.fretland@hist.no Find

Nr: Ty All types of tasks Tasktext: Author: Edit & add: Statistics: Delete:

1	Multiple-choice w/radio	Sikkerhet: Nei	(DM 2.1.14) List elementene i mengden $(A \cap B) - C$, når mengde-universet $U = \{1, 2, 3, \dots, 10\}$ og mengdene $A = \{1, 4, 7, 10\}$, $B = \{1, 2, 3, 4, 5\}$ og $C = \{2, 4, 6, 8\}$.	Truls Fretland	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Multiple-choice w/radio	Sikkerhet: Nei	(DM 2.1.29) I en gruppe på 191 studenter er det 10 som tar både fransk, handel og musikk; 36 tar både fransk og handel; 20 tar både fransk og musikk; 18 tar handel og musikk; 65 tar fransk; 76 tar handel; og 63 tar musikk. Hvor mange er det som ikke tar noen av fagene?	Truls Fretland	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Multiple-choice w/radio	Sikkerhet: Nei	(DM 2.1.62) Avgjør om påstanden er sann eller usann: $\overline{X - Y} = \overline{Y - X}$.	Truls Fretland	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Multiple-choice w/radio	Sikkerhet: Nei	List elementene i mengden $(A \cap B) - C$, når mengde-universet $U = \{1, 2, 3, \dots, 10\}$ og mengdene $A = \{1, 4, 7, 10\}$, $B = \{1, 2, 3, 4, 5\}$ og $C = \{2, 4, 6, 8\}$.	Truls Fretland	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Yes/No	Sikkerhet: Nei	Mengdene X og Y er delmengder av den universelle mengden U . Anta at mengde-universet for det kartesiske produktet er $U \times U$. Er følgende påstand sann? $\overline{X \times Y} = \overline{X} \times \overline{Y}$ for alle mengder X og Y .	Truls Fretland	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Improvement – *the difficult phase?*

- Calculations: from days to seconds
- Useful statistics at different levels
 - identification of (un)successful questions
- Pool + statistics = interesting searches
 - questions considered successful
 - frequently used questions (or never used)
 - poorly designed distractors
 - ... etc etc
- Automatic test generation based on certain criteria

The distractor problem

- Students should not be able to eliminate distractors
- Most teachers find the task of writing high quality distractors difficult...
 - ... but the system can provide automatic generation of distractors based on wrong answers from a similar free-text version
 - ... or from similar questions → templates

Where is Educa?

1. Bremen
- 2. Berlin**
3. Beijing
4. At school

Current use of the tool

- Campus and distance learners, mathematics and ICT courses
- Testing is used as part of the learning process and for assessment
- We are gaining experience
 - the question pool is increasing
 - the repository of test results is growing
 - feedback from teachers and students

*He who loves
practice without
theory is like the
sailor who
boards a ship
without a rudder
and compass
and never knows
where he may
cast.*

Leonardo da Vinci

Lessons learned

- The pool has proven successful
 - reuse
 - search for high-quality questions
- Statistics significantly help the teacher in the improvement phase
- Manual analysis of test results has helped to guide the implementation of useful statistics
- Students are very helpful resources for system development and identification of bugs

Future work

- The layout must be improved
- Some functionality is not yet implemented
- Explore possibilities for usage
- Analyse to which extent such a tool is efficient and a time-saver for the average teacher, and ensuring high quality
- Identify methods for best practices within online testing

Have a nice weekend

EASY

0	1
1	0

A LITTLE BIT
CHALLENGING

1	0
0	1

REALLY
DIFFICULT

1	0
0	1

IMPOSSIBLE

1	1

svend.horgen@hist.no

<http://www.aitel.hist.no/~svendah>