dutchTest

The Introduction of Large Scale Computer Adaptive Testing in Georgia

NUCEM conference

Bratislava, Slovakia

Wednesday 21 October

Two examples of CBT for National Exams

- The Netherlands
  - Started with computer-assisted testing
  - Long piloting phase (2000-2017) of CBT
  - Careful information and advocacy campaigns
  - Introduction only after full agreement by stakeholders on value-added
  - 2010: Dutch Board of Examinations commissioned development of integrated system to external provider
  - Delivery full linear tests over internet; same for all students but variants used during one-month testing window
  - No CBT across all subjects
- Georgia
  - Started from scratch in 2010
  - Nation-wide ‘pilot’ in May 2011 (44.000 students, 1500 test centres)
  - CAT for all subjects
  - Careful information and advocacy campaigns
  - Strong emphasis on involving schools; offering value added through customized feed-back

Dialectics of Progress?

Why CAT in Georgia? Political Context

MoES decided that

- school leaving exams had lost currency, and
- had to be replaced by external tests, which
  - should be administered locally;
  - should be secure beyond any doubt;
  - should not be a huge burden on the budget;
  - would put Georgia's on the list of high-tech knowledge economies.
- NAEC advised CAT:
  - Avoiding printing costs.
  - Making effective use of item banks.
  - Allowing for flexible continuous process with lesser demand on testing facilities
  - High level security, because each student would have his/her own test

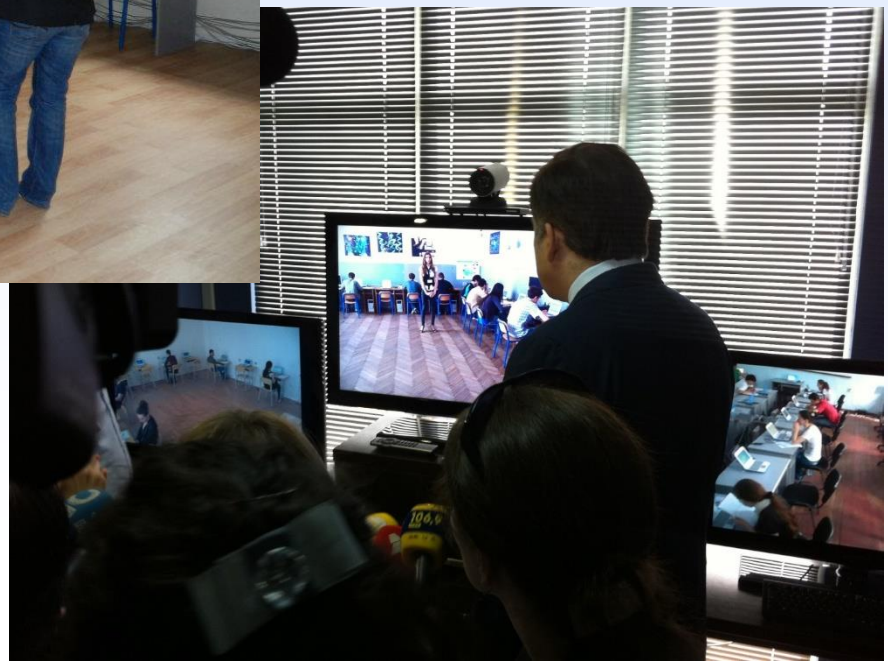
Planning the CAT; some figures

- Three psychometricians trained by CITO and US psychometricians.
- Additional training for test developers.
- 2300 proctors trained and certified by NAEC.
- 200 regional IT school support staff trained.
- Item banking software developed.
- CAT algorithms developed.
- Item banks for 8 subjects developed and calibrated.
- Servers and routers purchased for national centre.

Planning the CAT; some figures (ctd)

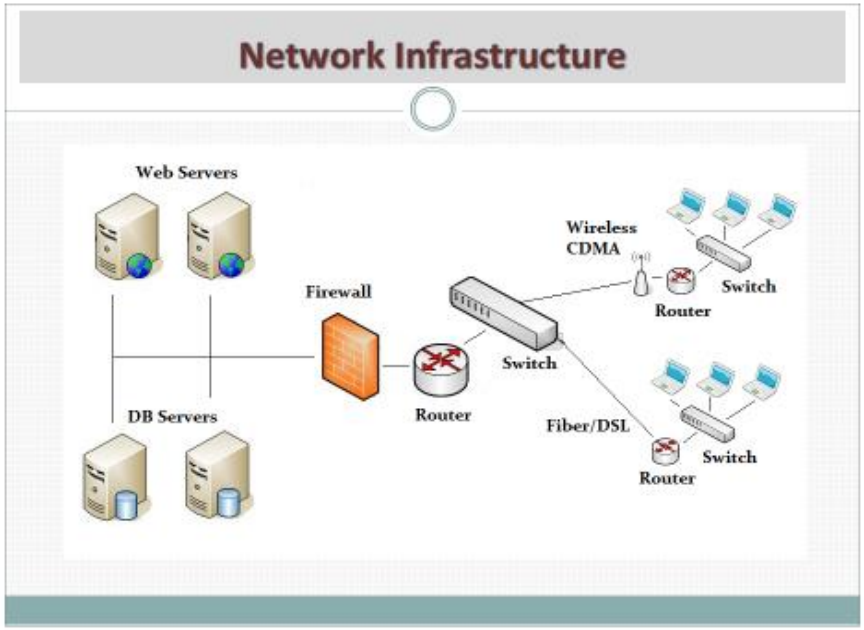
- 1800 surveillance cameras bought by NAEC for test centres.
- 11.000 computers purchased by MoES for use in the 1600 schools that were going to serve as a testing centre.
- 1600 testing centres to be connected to the internet (570 glass fibre, 1100 wireless connections).
- Twelve major regional information meetings held by NAEC; brochures and web-based practice tests prepared, Q&A on NAEC Facebook page, mock tests for all students (45.000) in all 1600 testing centres.

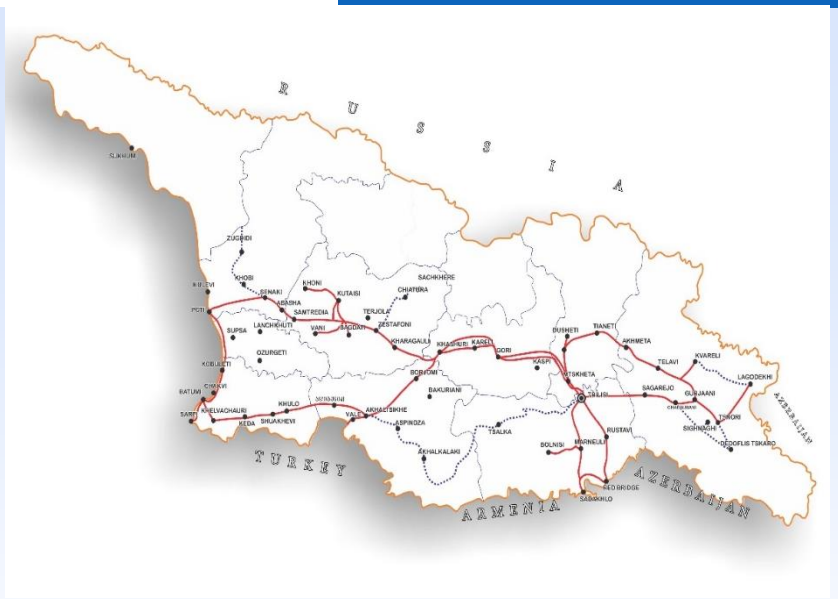
Testing Centres



dutchTest

Testing Centres: connectivity





570 schools fibre-optic 50 Mbps
1600 schools wireless; 3 Mbps down, 1 Mbps up
Student taking CAT needs 32 Kbps
Continuous buffering of all data of all logged-in students

Even Wi-Fi allows 30 students taking CAT at the same time

dutchTest

Testing Centres: security

- Main measure: well-trained and motivated proctors.
- Technical measures:
  - Software application installed on test centre computers preventing screen prints/dumps, copying of texts or graphics, use of external drives or other peripherals, and denying access to any site other than the NAEC CAT website.
  - Windows shell replacing the standard Windows interface and denying access to the standard applications running under Windows, connecting to web-based CAT application.
  - Firewalls and IP-filtering at central server
- Network providers: no interception of signals.
- Item leakage due to students memorizing them: few.

Administering the CAT

- Registration two month ahead of the testing
- Mock tests one month prior to the testing
- May 2011: 47.000 students
- May 2012: 45.000 students
- May 2013: National School Leaving exams cancelled
- Oct. 2013: Science CAT for grade 12
- May 2014: CAT for remaining subjects

dutchTest

The CAT's costs

- Many costs are 'hidden'.
- Main cost items:
  - Computers for testing centres; MoES invested a large sum for equipping the majority of Georgian schools with pc's, which also could be used in testing centres;
  - Surveillance cameras;
  - Item writers;
  - Test administration costs (registration, test centre management, NAEC office costs, transportation, accommodation and subsistence);
  - Proctors (the largest continuous cost item).

Appr. €1,92 M

dutchTest

Stakeholder Opinions

- School Principals
  - Positive; less fear of punitive measures and CAT makes their lives easier.
  - Appreciate feed-back provided by NAEC.
  - Regret that actual items cannot be seen; also for appealing .
- Teachers
  - Experience CAT as a fair, but limited means of assessment.
  - Regret that students cannot change answers once given.
  - Appreciate that CAT is not used for accountability purposes.

Stakeholder Opinions (ctd)

- Students
  - Positive; experience tests as fair, objective and not too difficult.
  - Concerns about validity: e.g. MFL speaking and writing skills not assessed.
  - Admit improvement in studying (all subjects; attending classes in grade 12).
  - Doubt decrease extra-curricular tutoring; instead see increase due to tutoring for SGE
- Media
  - Generally positive, also because pass rates were high.
  - Positive about technical and security aspects.
  - Some negative comments on limited validity.
  - Concerns about lenient cut scores covering up low competence level of Georgian students.

Success factors

1. Strong government commitment.
2. NAEC's leadership and stakeholders' confidence in NAEC's competence.
3. NAEC's strong psychometric and ICT competence.
4. NAEC's experience in large scale secure testing.
5. Smart test design avoiding network overloads and student data getting lost.
6. Full scale pretest under realistic conditions shortly before the real tests.

Caveats

- Doubts about the validity of the tests among stakeholders.
- Reliability of the ability estimates, both psychometrically and at face value.
- Security of items and right to appeal.
- Negative backwash effects caused by applying low cut scores.

dutchTest

Thank you!

http://siteresources.worldbank.org/INTREAD/Resources/Bakker_Introduction_to_CAT_Georgia_for_READ.pdf

http://go.worldbank.org/8D8GTBPLF0