E-Assessment in Mathematical Sciences — Newcastle University

Notes by J.H.Davenport

22 June 2020

Chapter 1

22 June 2020

1.1 Back to the Future: Matthew Hillier

Handwriting.

UNSW: 550k students, can have 2k in one class. Experiment with new Data Science course. Academics wanted a traditional clone. Paper not taken away, so no paper. All BYOD with online proctoring. 150 students in two cohorts. One staff member tried Microsoft Surface But writing on a Microsoft tablet means the camera is pointing in the ceiling, destroying the remote proctoring.

Experimented with an AUS$50 device (Huion?). Experimented with Microsoft, but failed. Looked ta Moodle. No handwriting, but looked at Poodll solution — already in our hosting services code base.

Upload speed is typically 1Mb in Australia (12 down 1 up) but proctoring provider wanted 3Mb up. 25% of comments after were Internet issues. Serious warning about remote proctoring companies. Note ETHZ uses Moodle for final exams: see his e-mail to me.

Chapter 2

23 June 2020

2.1 Assessing Mastery of Introductory Pure Maths: Jessica Banks (Liverpool)

Uses Mobius, but most ideas general. Taken by straight maths students on arrival Particular introduction to abstract pure. Topics include naïve set theory Note that we think students the core 40% to pass, but in fact they can pass on any 40%. Hence we require students to get *all* the assignments right *enevtually*.

Q Is this function injective? Piecewise linear, so each piece injective.

Q Several on notation, e.g. greek letters, $\exists$ etc.

Can set penalties, e.g. require two correct answers in a row. But the rules are pretty inflexible. Test available through weeks 6–7, resit in 11. Formative assignments (similar) before.

Q Mobile devices?

A Little experience, as Blackboard. Moving to Canvas, which has more moble support. No complaints here.

Comments about disengaged students, given the "Chinese drowning" experience.

2.2 LON-CAPA to Moodle Stack: Frauke Sprenkel

Maths/Theoreetucal CS/Programmings. Lon-capa has Jflap abd Graja for program grading. Quite complex. Wanted to continue Graja and JFlap (Grappa plus these into Moodle, our middleware). Lon-Capa had a large shared resource pool, but last release 2017, last major 2014.

The LaTeX, Maxima, images (SVG) all migrate, but there's no automated migration from Perl to Maxima.

1. Get permission
2. Get source code
3. Migrate
4. Test
5. Rewrite

Did 20 problems/week: quasi-automatic, significant rewrite/simplify/use the idea but write from scratch. Students like unity and look/feel but Stack is slower than Lon-Capa, and browser/OS issues. `github.com/kiliandangedorf`.

2.3 Success and challeges of Numbas: Mario Orsi (UWE)

High level of randomisation. Good shared facilities, team working. large database. Feedback incredibly positive. Replaces hgh-stakes exams. Slight increase in mean marks/%age pass. Attendance much improved, especially in tutorials. Exam no-shows when down from 4% to 1.7%.

But bug in Blackboard meant marks were often missing, and students had to grab screenshots and complain. Fixed, but UWE wouldn't import the fix as it involved IT work. Tryinf to get a contract with Mayne bio analytics.

2.4 Marking online: Edinburgh

Cancelled year 1/2, online 3/4/5. Time normal+1 hour. No detriment policy. Blackboard, so little control over what's submitted. Draft tested by postgrads, available in Moodle. Used Gradescope, RM Assessor[1], GradeX PDF (mostly small courses). 2400 scripts. Note CJS's first year course has 700.

RM Good workflow.

- Zoning (where is each question in the script) done by postgrads not marking 1/2 year.

Gradescope Allows retrospective rubric changes. Equivalent of zoning is hard.

GradeX PDF Very lowtech. Written in Go by an engineer, so I forked it. Adds typable forms to each page. Did the adding up, and helped exam boards locate borderline scripts etc.

Need to solve the zoning issue. Maybe "stack plus upload your working here".

[1] Used IB, SQA, English Boards.

2.5 Automarked first year: Chris Graham (Newcastle)

Used Numbas for whole of Stage 1 S2 in Maths, Stats, Physics. Numbas is deeply embedded in many subjects at Newcastle. 14 Numbas exams, e.g. Multivariable Calculus, Dynamics, also Marine Statistics (Natural Science). In Matsh we went for pass (60%)/fail with unlimited attempts in three-week period Score and mark breakdown after each attempt. In afct less than half passed first time.

Chapter 3

24 June 2020

3.1 JSXGraph

3.2 JSXGraph in Moodle: George Kinnear

Chapter 4

25 June 2020

4.1 NUMBAS Journey: Ann Smith (Huddersfield)

Had been putting together NUMBAS for some time (teaching in China) but it became obvious wouldn't go to China first, then it happened at home!. Knew that Youtube wasn't available in China. 40 credits worth in two blocks. `https://www.mathtutor.ac.uk`.

Real problem was actually level of English with some students. Numbas-5 alternative answers proved useful. Would I do it again: absolutely!

4.2 Turning to WobWork (Morris U, PA)

2 days to convert. Flooded with information. The students were already used WebWork in class minimally. I was multivariate calculus. Can use one-line iframe embedding to put my videos in the WebWork exercises. Feedback boxes let me communicate with the students who weren't reading e-mails. Also students could photograph their handwriting, to Google Drive, and put URL in essay box.

4.3 CUNY WebWork OER

CityTech (downtown Brooklyn). We could offer Webwork at zero cost (very important for us). Customisable is important.

Title V grant makes curation and language work important. Turnkey courses, especially for part-time Faculy, is really important. Up from 35 sections to over 100 using WebWork. CUNY central has added funding to share content across CUNY campuses (now 9 participating, out of 24).

Working forward, wor kon consistency of language is important.

Chapter 5

26 June 2020

5.1 Möbius

DigitalEd based in waterloo. 300+ schools, 300k students, 7M assessments graded automatically. Ecosystem to access and share content. Open Content, so you can take the bits you want, and adapt them. LaTeX with MathJAX: good for screen readers.

5.2 Shared Research Agenda on Computer-Aided Assessment: Kinnear

Basically an ongoing research project. But one can join.

5.3 Developer Update: Stack; Sangwin

Uses Maxima. Separates validity from correctness. Formative feedback in mind. New features include support for intervals, propositional logic, tables (e.g. truth tables). Thinking muc more about "proof comprehension sequences". Currently working on rpoducing STACK exercises for the HELM workbooks (+Loughborough etc., all to be released under CC). We *may* be able to host a Moodle server with HELM.

5.4 Developer Update: NUMBAS

Powers SCHOLAR for 400k students in Scottish schools. Now easy to embed in web pages without any scoring.

Vast amount of NUMBAS use in Maths at Newcastle is formative, so introduced menu mode as against exam mode.

Q Google translate

A Not good enough.

Michael Gage I like a range of 4–5 systems we can choose from.

5.5 Developer Update: Webwork

We can embed an iframe (currently webwork.rochester.edu) in any page.

Chapter 6

29 June 2020

6.1 Experiences with variable assessment data: Graham Clarke (RMIT)

Variable data reduces (doesn't eliminate) collusion. Sometimes randomly variable, sometimes numerically variable, sometimes symbolically variable.

Example 1 *Every student got a different Cayley table, but isomorphic. Inventing this was tedious.*

Example 2 *I used* $N := \sum_{digits}$ *student number (+1 is prime). Then asked to find groups of order* N*, or other questions.*

Example 3 (Error-Correcting Codes) *Use* N *as index into a table of syndromes for an ECC quesion.*

And many variants on this. But sometimes there were variations in difficulty.

Example 4 *Generate a six-character "special name" from family name, use in* `I AM XXXXXX FROM RMIT` *and do Huffmann coding etc.*

NB Some students got different results for "special name" than we expected.

Q Wrong parameter?

A We suspected people were deliberately getting the same as their friend, so were tough on this at exam board.

6.2 JHD

See `https://staff.bath.ac.uk/masjhd/Slides/AssessCoding.mp4`.

Q Why not submit via a version control repository.

A Good point about version control: when I started we didn't have a GitHub instance. I MIGHT do that, but the office staff are happy with the (University-mandated) Moodle/extensions process, especially with anonymous submission. Not sure how to interface that.

6.3 Inclusive e-Assessment: Christian Lawson-Perfect

There's a "deficit model" but it's better to think of all students from the beginning. Three issues: Subjectivity/Accessibility/Integrity[1]. See Hottinger's book "Inventing the Mathematician" for subjectivity.

NB

Q names?

A I use them based on distribution of names used for children born 18 years ago. Needs recomputing for different countries.

6.4 Online STEM Assessment: Rederly

Webwork-based apparently. Team of about 5. Also CUNY fellows, and the wider WebWork community.

6.5 Innovations in the E-Testing of Statistics for Non-Specialists: Ian Weir (UWE)

Level 1 module delivered to 400 business students at UWE. Challenging amount of material plus SPSS. Controlled: best 2 out of 3 of assessments set by Statistics. Also Uncontrolled, sets by Business. Dewis is UWE's open source e-assessment systems. There's a DEWIS-R interface: R generates the dataI just possed my R code into this. which goes to the students. These answers are analysed by R.

[1] "One cheat is less painful than 100 giving up".

Chapter 7

22 June 2020

7.1 Möbius at Birmingham

Jenang–Birmingham Joint Institute. Möbius as at Birmingham proper[1]. Y1 (2017/18) had 100 students. 80 credits used CAA for summative assessment. We had planned closed-book class test. I wanted assessment *for* learning, rather than *of* learning. 2019/20 we had 210/200/80 in years 1/2/3. So now 80×3 credits. Good slide on year-long cycle.

Jul-Aug Content creation

Sept Test

Oct–Dec run assessments; Nov. re-hire interns for S2 (mostly from summer)

Dec-Jan content creation

...

Interns code and test questions.

7.1.1 Principles

Developed over the

- Avoid MCQs where possible (avoids reverse engineering)
- Natural input (Maple and LaTeX are not LOs).
- Randomised versions, but need to ensure comparable difficulty, e.g. similar sparsity.

[1] Also STACK

7.1.2 Regrading

Have custom code to collect answers, and look at wrong answers. The commonest issue is syntax errors: Mobius doesnt distinguish syntax errors from wrong answers. Various recommendations. Internationalisation is a problem: East Asia has "thick brackets", and variants on commas. Typically 2–3 out of 20 in the first year need grading.

Note that Möbius will not allow to ask for a diagonalisation, since PDP^{-1} will get simplified away. Can't ask for P and D separately, as Möbius doesn't support linked boxes.

7.2 A New Moodle Question Type: Vernitski (Essex)

Knwon internally as Blackwater. Type's goal is to ask variants on a definition, e.g. changing quantifiers, swapping $<$ and $>$. Aim is to cut/paste from lecture notes, then add the variants. So has a syntax for indicating the variants.

Q Feedback — I like the style "if you said this, you probably made this mistake".

A Not really.

Q How to use?

A Happy to share.

Chapter 8

1 July 2020

8.1 Studentstart.de: online student pre-course in Mathematics for Engineers

Baden–Württenburg Cooperative State University. 6000 students in Mannheim.

1. Diagnostic pre-test
2. Learning suggestion
3. Self-study + e-advice
4. e-tutorials
5. September/November face-to-face

"Course 0", six courses covering school material, and five more specialist post-test.

Ilias as the LMS.

Example 5 (Functions and roots: lesson 4) *Has Goegebra examples. Fedback such as "you have x and y swapped".*

Some courses use STACK. Typically pre-test 50%, post-test 55%. But it was the weaker students who improved more. Hence role of formative assessment.

8.2 Stack Project at Southampton

First year Caculus e-assesment final DEWIS/STACK. Aim to avoid "feedback graveyard". So use e-assessment for computational questions, but human marking immediately 1:1 with student for "proof" questions. Lovely quote from Babbage.

8.3 Japanese STACK

JHD had to leave at this point.