Transforming Exams Across Australia
Australian Government Office for Learning and Teaching
National Grant ID15-4747: AU$500K over 3 Years

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr Mathew Hillier;</td>
<td>University of Queensland</td>
</tr>
<tr>
<td>Leader OLT national project. Seed leader</td>
<td>Monash University 5 October 2015</td>
</tr>
<tr>
<td>Dr Andrew Fluck;</td>
<td>University of Tasmania</td>
</tr>
<tr>
<td>Originator of USB e-exam concept. Seed partner.</td>
<td></td>
</tr>
<tr>
<td>Dr Michael Cowling, Mr Kenneth Howah</td>
<td>Central Queensland University</td>
</tr>
<tr>
<td>Seed trial site.</td>
<td></td>
</tr>
<tr>
<td>Dr Kim Blackmore</td>
<td>Australian National University</td>
</tr>
<tr>
<td>Assoc. Prof. Paul Newhouse</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>Dr Matthew Bower, Prof. Dominic Verity</td>
<td>Macquarie University</td>
</tr>
<tr>
<td>Prof. Marilyn Baird (H), Mr Scott Grant (A)</td>
<td>Monash University</td>
</tr>
<tr>
<td>Assoc. Prof. Shona Leitch</td>
<td>RMIT University</td>
</tr>
<tr>
<td>Dr Ruth Geer, Mr Bruce White</td>
<td>University of South Australia</td>
</tr>
</tbody>
</table>

TransformingExams.com

Acknowledgement: Support for this project has been provided by the Australian Government Office for Learning and Teaching. The views expressed do not necessarily reflect the views of the Australian Government Office for Learning and Teaching or participating institutions.
Targeting...

- Supervised
- High stakes
- On campus
- Large scale

*(image credit: Dr Fluck UTAS)*

What we are **not** specifically addressing here is off campus, online only, distance education, cross institutional students – there are extra issues (later!) and some possible e-solutions to address these needs.
National Project Phases

2015 Start-up, staffing etc (Move to Monash Uni Oct).
• First briefing session held 13 Aug 2015 (recording available).
• Next steps:
  – Prepare ethics protocols at each partner site.
  – Prepare business cases/briefing notes at each partner site.
  – Set-up project document repository (AARNET in progress)

2016 Software development/upgrades
• Upgrade to version 14 or 16 of Ubuntu OS.
• Continuation of trials at original partners (UQ, CQU, UTAS) and ground work at Monash.)

2017-2018 National trials at new partners.
• All nine sites running at least one trial. (others are welcome)
• Continuation of tech development.

2018 (end) Final report.
**Lead institution***. Funded. 'e-Exam HQ': Project leader > project manager, technical development staff, lead RA. Trial site. Mentor.

Funded partner institution(s) x 2. [UTAS, CQU]. Run multiple trials. Collect/return data, provide mentorship. Staff: Site lead and casual RA.

Self-funded partner institutions running trials x 6 [ANU, ECU, Macquarie, Monash*, RMIT, UniSA]
Receive mentorship; collect/return data for min 1 trial in years 2 or 3. Staff: Site lead.

**External Evaluator**

- a) International technology exchange partners e.g. Finland Govt, other Unis.
- b) Other universities interested in running trials (self funded).

Other institutions welcome!

*Lead*: Currently UQ but Monash will take on lead after Oct 2015 after my move.
Start-up Processes for Partners

Process matters – for project partners.

• HQ will provide a trial 'pack' of documents incl ethics protocol template, instruction manuals, surveys etc.

• Preparing for your trial - get a budget for hardware (USBs & copy box) and staffing (IT helpers, RA). HQ can provide advice. Suggest you source internal grants and/or direct requests.

• Preparing your brief / business case for internal approvals.

• We are establishing online work area via AARNET
  – Cloudstor project document repository – all project partners will have access soon (passworded)
  – Public website [http://transformingexams.com](http://transformingexams.com)

• Other institutions welcome! ... Ako Aotearoa $?
e-Exams Seed Work 2013 to 2015
OP NZ Aug 2015

OLT Project leader / Presenter: Dr Mathew Hillier, University of Queensland
OLT Project collaborator: Dr Andrew Fluck, University of Tasmania
OLT Project system developer: Marisa Emerson, University of Queensland
UQ course academics:
Dr Arosha Weerakoon (Dentistry)
Dr David Booth (Zoology),
Elizabeth Springfield (Occupational Therapy),
Katrina Williams (Physiotherapy),
Prof. Malcolm Jones (Veterinary Biology),
Rebekah Scotney (Veterinary Technology) and
Dr Robin Fitzgerald (Criminology)

Get the demo and user guides
http://transformingexams.com

Acknowledgement: Support for this project has been provided by the Australian Government Office for Learning and Teaching. The views expressed do not necessarily reflect the views of the Australian Government Office for Learning and Teaching or participating institutions.
Is this your exam space?
Targeting...

- Supervised
- High stakes
- On campus
- Large scale

(image credit: Dr Fluck UTAS)

What we are not specifically addressing here is off campus, online only, distance education, cross institutional students – there are extra issues (later!) and some possible e-solutions to address these needs.
Rationale: Concerns, drivers, possible solutions for e-Exams. A truly 'wicked', 'messy' problem and a long road to get it right!...

## e-Exams: Online, Offline, On Campus or Distance

There are trade-offs for any e-exam solution.

<table>
<thead>
<tr>
<th></th>
<th>Online</th>
<th>Off-Campus</th>
<th>On Campus</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Space issues for institutions.</td>
<td>• No space issue for institutions.</td>
<td>• More efficient exam management.</td>
<td>• Less secure: students at home.</td>
</tr>
<tr>
<td></td>
<td>• Improved exam management efficiency.</td>
<td>• More efficient exam management.</td>
<td>• Students supply equipment.</td>
<td>• Needs reliable network.</td>
</tr>
<tr>
<td></td>
<td>• Equipment: computer labs big enough to cater for 2000 at once.</td>
<td>• Less secure: students at home.</td>
<td>• Students supply equipment.</td>
<td>• Network reliability not an issue.</td>
</tr>
<tr>
<td></td>
<td>• More secure: it is supervised.</td>
<td>• More secure: it is supervised.</td>
<td>• Needs reliable network.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Needs reliable network.</td>
<td>• Network reliability not an issue.</td>
<td>• Network reliability not an issue.</td>
<td></td>
</tr>
</tbody>
</table>

There are trade-offs for any e-exam solution.
The core issue! ...

We are faced with a growing disconnect between the way *high stakes testing* is conducted using pen on paper exams and students’ everyday experiences of study and life.
Where we are going: Post-paper exams

We need greater pedagogical flexibility and more authentic assessments in the exam room. ... alignment!

Simulations, tools of the trade, virtual experiments...

'Manuals' software via WINE. E.g. CAD / 3D modeling, Celestia.

Moodle quiz with media (auto marked).
Students type in a short sentence response which can be marked by computer based on pattern matching.

Available in Moodle now.

Example question

A boy climbs slowly to the top of a slide and then slides down it. At which point will his kinetic energy be a maximum?

Note: Your answer should ignore the effects of friction.

You should give your answer as a **short** phrase or sentence.

Kinetic energy will be at maximum when at the bottom of the slide.

Sally Jordan, Open Uni UK, TA webinar, 5 June 2013 transformingassessment.com/events_5_june_2013.php
Examples – Confidence questions

Confidence based approaches penalise guessing. Students need to choose a response and declare their level of certainty. Available in Moodle now.

Certainty levels and consequences

<table>
<thead>
<tr>
<th>Certainty level:</th>
<th>C=1</th>
<th>C=2</th>
<th>C=3</th>
<th>No Reply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark if correct:</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Penalty if wrong (T/F Q)</td>
<td>0</td>
<td>-2</td>
<td>-6</td>
<td>0</td>
</tr>
</tbody>
</table>

Qu. 1:
Which structure shown below represents meso 2,3-dichlorobutane, A, B or C?
[Click on the text below to open a window with the three choices]

Three structures, A, B and C

Choose one of the following:
- C
- A
- B

No Reply  | Certainty:  | C=1 (low)  | C=2 (mid) | C=3 (high) |

University College London

Tony Gardner-Medwin, UCL, TA Webinar 6 April 2011
# e-Exam System Affordances

<table>
<thead>
<tr>
<th>Pertinent Features</th>
<th>Affordances</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 'Whole computer' environment (OS, LMS, applications...) on a stick.</td>
<td>Vastly expanded pedagogical scope over that of a browser window.</td>
</tr>
<tr>
<td>Typed student responses via Word processor, constructed via apps (human marked) or on-board learning management system quiz (computer marked).</td>
<td>Caters for introduction to advanced uses. Components added/removed to suit. Electronic collection facilitates analytics, item response analysis...</td>
</tr>
<tr>
<td>No live network required during exam, even for LMS questions.</td>
<td>Robust. Greater control. (network could be used for admin)</td>
</tr>
<tr>
<td>Student owned equipment used as host and left untouched.</td>
<td>An ethical approach to scalability (no invasive software to install)</td>
</tr>
<tr>
<td>Modular, open source code base and commodity 'off the shelf' components.</td>
<td>Leveraging popular and sustainable projects for better efficiency. Fully 'known' (no 'blackbox'). Available!</td>
</tr>
<tr>
<td>One version works on most Intel based laptops - Apple, 'windows', Linux, that have a USB port.</td>
<td>One software version can serve all. Streamlines development and maintenance.</td>
</tr>
</tbody>
</table>
Where we are now: Paper Equivalent


**Question 2.** Match the following host-MOTA and host-TOTA (marked in red below).

Possible descriptions:
- a) Mauris id mi id orci interdum semper.
- b) Sed eu neque ut est dignissim fringilla.
- c) Vivamus in dolor euismod, luctus libero.
- d) Mauris vehicula eros a viverra pellentesque.
- e) Curabitur eu mi at nibh commodo varius.
- f) Acenae eget orci porta, malesuada lorea.

Please write or type the letter of the descriptions listed above.

<table>
<thead>
<tr>
<th>Answer a to f</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>I. Paxogen</td>
</tr>
<tr>
<td>a</td>
<td>II. Sitabosis</td>
</tr>
<tr>
<td>c</td>
<td>III. Fakecasalism</td>
</tr>
<tr>
<td>e</td>
<td>IX. Benzol</td>
</tr>
<tr>
<td>g</td>
<td>X. Willion</td>
</tr>
</tbody>
</table>

**Question 5:** For the following diagram please provide the names for THE XING in the table below.

| A | Label goes here. Constructed response question. |
| B | Blue text makes it easier to see which questions have been answered and which have not! |
| C | Use minimum row heights to provide plenty of space, but don't use double carriage returns! |
| D | Doing so means the layout is less likely to be disrupted. |

**Question 3.** Samuel is 5 years old and attends racing cars 5 days per week. Eamon is 10 years old and rides a superbike around the same track. It is not a selected response item so some text will be expected. In the table below, give two (2) examples of flippant laxism relevant to his age range (4-6 years), and describe how Samuel and Eamon differ in their abilities to perform laxism.

<table>
<thead>
<tr>
<th>Term</th>
<th>Control</th>
<th>A &amp; B</th>
<th>Samuel’s abilities (age 5)</th>
<th>Eamon’s abilities (age 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>[4 marks]</td>
<td></td>
</tr>
</tbody>
</table>

**Question 7:** Some rationales for punishment are XEZA TTT. What does this mean?

The student types their answer here. In this example, a two row table. The response table row is created and a minimum height cell instead of successive carriage returns to set the box height, the next question will be less likely to be disrupted when students type their responses. The initial size of the box should indicate the desired length of the response. The box will automatically expand when it gets full.

Please write / type your response inside the box below.

More details about setting heights appear later in these examples.
e-Exam Trials Workflow

Set-up: prepare exam learning materials

- Academic creates exam learning material
- Create master USB (tested)
- USBs duplicated per student

Pre-session: Student laptop setup & practice.

- e-Exam system takes over laptop. Ubuntu Live USB. Libre Office.

Exam room use

Post session: retrieve responses and assessment

- Collect USBs (responses)
- Responses retrieved from USBs.
- Collated e-responses sent to academic.

1. Students enter room.
2. Given USB.
4. Do exam.
5. Return USB.
First and Most Recent e-Exams

VETS2100 S2 2014

Used standard teaching rooms, sought rooms with tables and power sockets.

DENT4092 S1 2015

← VETS: hand-writers sat in rows.
Attempted to separate typists and hand-writers where possible.
DENT: typists at the back, → hand-writers at the front.
Floor Plans

Purple = typists
Orange = hand-writers
Walk Through

1. Start with the computer turned OFF. Then insert USB stick

2. Hold down ALT then power on.

   Apple
   Press and release power
   Hold down
   Keep holding ALT until you see...

2. Power on while tapping ‘boot key’ (e.g. F12 or... )

   Other/Windows*
   Tap tap tap...
   Press and release power
   Keep tapping until you see...(or similar)

*Win 8: first need to disable secure boot.
Walk Through

Apple
3. Select a yellow icon.
EFI boot or ‘Windows’

Could be either one! So just try.
If you get
Try the other one!

Other/Windows
3. Select USB device.

It might be labeled something else and probably won’t be first.
Walk Through

4A. Some system messages may appear, if so just wait and see.

```
ata_id[292]: NDIO_GET_IDENTITY failed for '/dev/sdb': Invalid argument
```

4B. Exam system should start.

5. Arrive at e-Exam system desktop.
Walk Through

6. Student now types in their student ID number and name. Click Start Exam.
Walk Through

7. Exam file opens ready to enter exam details and responses ('Word Doc' shown).

Note: original file copied and student number prefixed to file name.
8. Student types responses into areas indicated. (Note: Inserting ‘links’ into the question text is possible that can lead to multimedia or other software tools for constructed responses).
Walk Through

10. Shut down the system. When the system has powered off, hand in the USB Stick.
Data collected from students

• 2013 pre-project online survey (UQ wide: brief results shown)
  – UQ students surveyed about their preconceptions about e-exams.
  – ASCILITE paper Hillier 2014

• 2014 Trials pre-exam short survey (six courses – result not shown).
  – Conducted at the pre-exam practice setup sessions.
  – Covered: student preliminary impressions, technical hardware compatibility.

• 2014 Trials post-exam extended survey (six courses – results shown)
  – Conducted at the conclusion of the exam (in the room).
  – Covered: rationale, student exam experience, reaction to exam session conditions, e-exam system impressions, exam writing strategies and production, general non-exam writing strategies.

• 2015 Analysis of text production (new trials; 1 course results shown)
  – Marks v word count, typing v handwriting, more to come (language density...)

UQ e-Exam Trial Outcomes
DENT Mid Semester Exam S1 2015 – Six short answer questions; 20% of course.

Scripts N = 68 (19 typed, 49 handwritten)

**Score**
Mann-Whitney
U 295,
Z -2.333,
Sig .02

**Words**
Mann-Whitney
U 245.5
Z -3.007
Sig .003
Comparing number of words typed and handwritten by question number.

Significance by question Mann-Whitney:
Q1 > .05
Q2 > .05
Q3 NS
Q4 > .05
Q5 > .05 close
Q6 NS

Typists produced more text.
Does more words mean better marks?

DENT exam.

It depends!

A higher hand written word count generally led to passing and better marks, but lesser words did not always result in poor marks.

Typists did better overall. More typed words only slightly increased marks.

_Not claiming causation!_
More information....
Demo set-up Guide,
Student Practice and User Guide
http://transformingexams.com

Demo videos start-up, use and recovery examples.
Apple http://ta.vu/eexam-demo-a
'Wintel' (Dell) http://ta.vu/eexam-demo-d
Contact: m.hillier[at]uq.edu.au
References


220 more at: https://www.zotero.org/groups/e-assessment/items/tag/e-exam
1. Monday, 8 September - GMT 07:00: **BYOD on-campus e-exams at University of Tasmania (UTAS)**. Andrew Fluck, University of Tasmania, Australia.

2. Wednesday, 10 September - GMT 07:00: **Bring-your-own-laptop e-exam for a large class at NUS**. Seow Teck Keong and Jeffery Tay, National University of Singapore.

3. Thursday, 11 September - GMT 07:00: **Large scale fully online BYOD final exams: Not your parents multiple choice**. Rob Peregoodoff, University of British Columbia, Canada.

4. Friday, 12 September - GMT 07:00: **Finland's national matriculation exams goes electronic**. Matti Lattu, Matriculation Examination Board, Finland.

5. Tuesday, 16 September - GMT 07:00: **eOSCE - robust real time electronic marking for clinical examinations**. Sebastian Hunkeler and Dr Philippe Zimmermann, Institute of Medical Education, University of Berne, Switzerland.

6. Wednesday, 17 September - GMT 07:00: **Gamification of Clickers with BYOD**. Paul Lam, Chinese University of Hong Kong.

7. Thursday, 18 September - GMT 07:00: **Safe Exam Browser: A modular approach to secure and flexible online-exams**. Daniel R. Schneider and Tobias Halbherr, Swiss federal Institute of Technology Zurich.

8. Friday, 19 September - GMT 07:00: **Ten Years of e-Exams at Freie Universitat Berlin: an Overview**. Alexander Schulz & Nicolas Apostolopoulos, Free University Berlin, Germany.
Cite this resource
Results from 2014 trials
(not shown at the conference)
## Research Study Phases

<table>
<thead>
<tr>
<th>Phase 1</th>
<th>Institution wide online pre-project survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 2, Step 1</td>
<td>e-Exam Trial Expression of interest</td>
</tr>
<tr>
<td>Phase 2, Step 2</td>
<td>Pre-exam preparation survey</td>
</tr>
<tr>
<td>Phase 2, Step 3</td>
<td>Type the exam</td>
</tr>
<tr>
<td>Phase 2, Step 4</td>
<td>Handwrite the exam</td>
</tr>
<tr>
<td>Phase 2, Step 4</td>
<td>Post-exam survey</td>
</tr>
</tbody>
</table>

### Participation in Phase 1: approx. 928 respondents (Nov 2013 - Nov 2014)

### Participation in Phase 2: Six courses in 2014, two in 2015

<table>
<thead>
<tr>
<th>Steps of trial</th>
<th>Yes will type</th>
<th>Maybe type</th>
<th>Total typists</th>
<th>Attrition</th>
<th>No (hand-write)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 EOI</td>
<td>201</td>
<td></td>
<td>201</td>
<td></td>
<td>361</td>
</tr>
<tr>
<td>2.1 Pre - before try</td>
<td>94</td>
<td>16</td>
<td>110</td>
<td>91</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Pre - after try</td>
<td>86</td>
<td>15</td>
<td>101</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>4 Exam (after)</td>
<td>71</td>
<td></td>
<td>71</td>
<td>30</td>
<td>450</td>
</tr>
</tbody>
</table>
Phase 1 Survey Design

• Pre-project Survey constructed to cover a range of possible concerns.


• A survey by Dermo (2009) provided the core.
  – Acknowledge that we would be using it in a different manner (‘pre’ rather then ‘post’).
  – But! we never intended to replicate it, instead we used this as a means for eliciting student concerns across a range of issues.
  – Students would largely be responding speculatively based on their *preconceptions*. (instructions given to students accordingly)
Mobile device ownership (excluding desktop computers) was an average of 2.3 devices per student (standard deviation of 0.8).
## Phase 1 The questions

<table>
<thead>
<tr>
<th>Theme</th>
<th>Five point Likert scale from 1 “strongly disagree” to 5 “strongly agree”</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective factors</td>
<td>Using a computer for an exam is more stressful than a handwritten paper exam</td>
<td>2.9</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>I am at a disadvantage when undertaking computerised exams</td>
<td>2.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Teaching and learning</td>
<td>Computerised exams are consistent with contemporary learning approaches at university</td>
<td>3.8</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>The potential for immediate feedback with a computer based exam could help improve my learning</td>
<td>4.0</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Computerised exams allow me to demonstrate my knowledge in more ways than paper based exams</td>
<td>3.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Validity</td>
<td>Computerised exams are appropriate for my discipline/subject area</td>
<td>3.4</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Computerised exams need to include a variety of question types in order to test my knowledge fully</td>
<td>3.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Reliability</td>
<td>The technology used in computerised exams is unreliable</td>
<td>3.0</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Computerised exams favour some students more than others</td>
<td>3.5</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Paper-based exams are fairer than computerised exams</td>
<td>3.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Practicality</td>
<td>Technical problems make doing exams via computer impractical</td>
<td>3.3</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Doing exams in the campus computer labs is impractical</td>
<td>3.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Security</td>
<td>Computerised exams are just as secure as paper-based exams</td>
<td>3.3</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>It is easier to cheat in computerised exams than with paper-based exams</td>
<td>3.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Production</td>
<td>I prefer typing rather than hand writing essay answers</td>
<td>3.8</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>I work more effectively when I type on a familiar keyboard</td>
<td>4.1</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>I would prefer to use my own laptop to undertake a computerised exam rather than use equipment supplied by the university</td>
<td>3.7</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>I get hand cramps when handwriting exams of 1.5 hours or more</td>
<td>3.7</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>I would like to be able to type answers in an exam</td>
<td>3.3</td>
<td>1.4</td>
</tr>
<tr>
<td>Adoption</td>
<td>I want computerised exams replace paper-based exams at university</td>
<td>2.8</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Plus two open ended comment questions
Phase 1 Participation

• 488# students (37% males, 63% females) = 1%*
• 9% post-grad, remainder were undergrads (with an even spread across year levels).
• 45 programs: those with at least 10 are listed:

<table>
<thead>
<tr>
<th>Program</th>
<th>N</th>
<th>Program</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied science</td>
<td>25</td>
<td>Electrical engineering</td>
<td>13</td>
</tr>
<tr>
<td>Arts</td>
<td>60</td>
<td>Information technology</td>
<td>15</td>
</tr>
<tr>
<td>Biomedical science</td>
<td>24</td>
<td>Law</td>
<td>29</td>
</tr>
<tr>
<td>Business management</td>
<td>24</td>
<td>Mechanical engineering</td>
<td>25</td>
</tr>
<tr>
<td>Chemical engineering</td>
<td>11</td>
<td>Mechatronic engineering</td>
<td>13</td>
</tr>
<tr>
<td>Civil engineering</td>
<td>18</td>
<td>Pharmacy</td>
<td>16</td>
</tr>
<tr>
<td>Commerce</td>
<td>22</td>
<td>Psychological sciences</td>
<td>15</td>
</tr>
<tr>
<td>Computational mathematics and physics</td>
<td>13</td>
<td>Social sciences</td>
<td>10</td>
</tr>
<tr>
<td>Education</td>
<td>11</td>
<td>Software engineering</td>
<td>10</td>
</tr>
</tbody>
</table>

*Krejcie & Morgan (1970) state that for a population of 50,000 a random sample 381 would be sufficient to be representative in relation to opinions expressed by respondents to 95% confidence. (it wasn’t random).
#analysis performed on responses received at Feb 2014. Subsequent responses up to Nov 2014 were 928.
a. Themes drawn from open ended questions on currently held ‘concerns’ about e-exams and general comments. **Findings next slides.**

b. Statistics* used to explore the *body of opinion* represented by Likert scales rather than as a search for a single truth. Tended to stick to non-parametric tests.

Phase 1 Themes

Preconceptions.

The technology being unreliable stresses me out more than the thought of doing the exam – *Law student*.

A real programmer would be looking up the APIs for their language every time they wanted to do something, but they can't because they're forced to only use paper-based notes they have on hand. It's infuriating - *Computer science student*.

It's true that 'computerised exams favour some students more than others' - i.e., the ones that are proficient typists over the ones that aren't - but the same is true of paper-based examinations, which favour those with the ability to work through strong pain in their writing hand – *Arts student*.

Emergent themes from Phase 1 survey

- Technical reliability
- Cheating
- Match with discipline
- Keyboarding prowess
- Computer literacy

Comment Count

As a mature aged student, I would feel at a disadvantage doing a computerised exam as I am not as computer literate as many of the younger students - *Chemistry student*.
Phase 1 Findings: Match to Discipline

e-Exams suit my discipline - By program (major)

See Hillier (2014) for more.
Phase 1 Findings: I would like to Type

“I would like to be able to type answers in an exam” By program (major)

See Hillier (2014) for more.
Phase 1 Interim Conclusion

Students were

• Cautiously optimistic
• Just over half would like to see an e-exam option. Mean agreement (3.3) “I would like to be able to type answers in an exam.”

• Were attune to the nature of their discipline and how the idea of an e-exam might fit.
• The fear of the unknown (?) esp regarding technical failures and reliability.
UQ e-Exam Trials 2014; 2015

Data collected from students (opt-in S1 & S2 2014)

• Via pre-project online survey (UQ wide):

• Via pre-exam short survey (typists only - six course results next).
  – Conducted at the pre-exam practice setup sessions.
  – Covered: student preliminary impressions, technical hardware compatibility.

• Via post-exam extended survey (six course results shown later)
  – Conducted at the conclusion of the exam (in the room).
  – Covered: rationale, student exam experience, reaction to exam session conditions, e-exam system impressions, exam writing strategies and production, general non-exam writing strategies.

• Analysis of text production (2 new trials; 1 course analysed)
  Marks v word count, typing v handwriting, more to come (language density...)

47
Impressions of the e-exam system

Data collected from students (S1 & S2 2014) at pre-exam set-up/practice sessions.
Trial Phase: Pre-exam Survey

Data collected from students (S1 & S2 2014) at pre-exam set-up/practice sessions.

Before trying e-exam

After trying e-exam
Trial Phase: Pre-exam Survey

Open text comments – concerns and praise

- Fear of technical failure/crash
- Remember procedure/use difficulties
- Scrolling/touchpad use
- Power/battery
- Fear of data loss
- Drawing/diagram difficulty
- Fear of computer damage
- Newness/unfamiliarity
- Security
- Panic
- Eye strain
- Fail the exam
- Rules/contingency
- Good general positive/praise
- Good ease to use
- Convenient
- Fence sitting

- That my computer may freeze or not work properly.
- Random technological malfunction.
- Simply a different format of exam. Remembering this for the exam.
- Am I allowed to use a mouse? Will it affect the system? Scrolling opposite.
- Would it save my answers properly.
- Availability of power. Battery run out.
- Doing something wrong and panicking.
- Good concept. Provides an alternative to written exams. Easy to navigate.
- Fairly sure I'll use the laptop but just need to think about it a little more.
Pre-exam Laptop Testing

Number of laptops passed and technical issues.

So... we had a pool of loan laptops.

Upgrade to next version of OS will help too.
Data collected from students (opt-in S1 & S2 2014)

• Via pre-exam project online survey (UQ wide):

• Via pre-exam short survey (typists only - six course results).
  – Conducted at the pre-exam practice setup sessions.
  – Covered: student preliminary impressions, technical hardware compatibility.

• Via post-exam extended survey (six course results shown next)
  – Conducted at the conclusion of the exam (in the room).
  – Covered: rationale, student exam experience, reaction to exam session conditions, e-exam system impressions, exam writing strategies and production, general non-exam writing strategies.

• Analysis of text production (2015: 2 new trials; 1 course analysed)
  Marks v word count, typing v handwriting, more to come (language density...)

UQ e-Exam Trials 2014
### Participation across the six courses in the 2014 trials

<table>
<thead>
<tr>
<th>Course</th>
<th>Description</th>
<th>Typed</th>
<th>Handwrote</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANIM (Animal Biology)</td>
<td>45 min mixed short answer and MCQ (type 'x')</td>
<td>5</td>
<td>109</td>
</tr>
<tr>
<td>BIOL (Zoology)</td>
<td>50 min short answer (Multiple choice section done pen on OMR sheet)</td>
<td>10</td>
<td>81</td>
</tr>
<tr>
<td>CRIM (Criminology)</td>
<td>70 minutes. Single long essay response section (and a Multiple choice section done pen on OMR sheet)</td>
<td>17</td>
<td>50</td>
</tr>
<tr>
<td>OCTY (Occupational Therapy)</td>
<td>100 min mixed short answer and MCQ (type 'x')</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>PHTY (Physiotherapy)</td>
<td>15 min (watch video and write) before OSCE</td>
<td>25</td>
<td>108</td>
</tr>
<tr>
<td>VETS (Veterinary Technology)</td>
<td>90 min theory, mostly short answer</td>
<td>11</td>
<td>78</td>
</tr>
</tbody>
</table>

It is important to note:
- First ‘toe in the water’ trials.
- Participation was optional.
- Mid term exams worth an average of 15% of the course grade.

**Gender**

<table>
<thead>
<tr>
<th>Typists</th>
<th>Hand writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>27 (40%)</td>
</tr>
<tr>
<td>Female</td>
<td>41 (60%)</td>
</tr>
</tbody>
</table>

Detailed case descriptions available: [http://transformingexams.com/uq_trials/UQ_e-exam_cases_s1_and_s2_2014.pdf](http://transformingexams.com/uq_trials/UQ_e-exam_cases_s1_and_s2_2014.pdf)
Reasons for typing the exam

(added 30 October 2014)

More time because of good typing skills
Poor hand writing skills, legibility for examiners
Editing potential
I think best when I type
Don’t get writer’s cramp
Saves paper
Prefer a screen
Phase 2 Post-exam Typist’s Comments

Quicker typing and the ability to edit or completely delete my answer without compromising on space.

I have terrible handwriting. Felt bad about it.

It is cleaner, I make lots of mistakes when I'm writing and it usually ends in lots of scribbles everywhere.

You can write as much as you otherwise would but don't get a sore hand when typing.

I could get info down faster and examiner could read it.

Typing is more natural for me. I think best when I am typing.
Reasons for handwriting the exam

(added 30 October 2014)

Computer issues/Fear of technology failure
Prefer handwriting, familiarity
Poor typing skills
Connect and collect my thoughts using handwriting
Handwriting is faster
Typing is more stressful
The noise from typing
Need to draw diagrams and scribble
Actual problems typing exam
No laptop
Don’t want to bring laptop to university
Formatting answer on paper
Not registering for session
Don’t like working on screens
Want a choice
Lack of experience
Previous poor experiences
Forgot to attend session
Health issues using screens
Phase 2 Post-exam Hand-writers

I felt more comfortable handwriting as nothing can go wrong & I wasn't relying on the computer system to complete my exam.

Three years of prior exams writing so stick with what you know.

I am worried about computer malfunctions.

I'm a slow typer and feel disadvantaged.

I was initially planning to type this exam but decided against it due to the unpredictability of machines.

I think more about what I'm writing when I handwrite but my hand gets sore and it isn't fast.

It's easier to handwrite. Though probably not easier for you to read my writing.

Lazy to bring laptop.
Likert scale/rating: 1 = strongly disagree to 5 = strongly agree [N = 69]. Means shown.
Did typers think the exam suited the use of computers?

Those that typed the exam.
All six cohorts combined (ANIM, BIOL, CRIM, OCTY, PHTY & VETS).
Likert Scale: 5 = Strongly Agree, 1 = Strongly Disagree

Mean of 4.2 (value shown) N = 69.
Largely that was a ‘yes’.

Those that typed the exam by cohort:
Student reaction to exam conditions

Typers (left) and Hand writers (right)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Typers Mean</th>
<th>Hand writers Mean</th>
<th>U</th>
<th>Z</th>
<th>Sig. (2-tail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall my experience of this exam was positive</td>
<td>4.04</td>
<td>3.48</td>
<td>13242.5</td>
<td>-2.132</td>
<td>0.033</td>
</tr>
<tr>
<td>I ran out of time</td>
<td>3.76</td>
<td>3.49</td>
<td>15203</td>
<td>-0.083</td>
<td>0.934</td>
</tr>
<tr>
<td>I felt more stressed in this exam than I normally do in other exams</td>
<td>2.61</td>
<td>2.69</td>
<td>14527.5</td>
<td>-0.751</td>
<td>0.452</td>
</tr>
<tr>
<td>I went back and read over my responses before submitting</td>
<td>2.61</td>
<td>2.56</td>
<td>15145.5</td>
<td>-0.394</td>
<td>0.694</td>
</tr>
</tbody>
</table>

Likert scale: 5 = strongly agree, 1 = strongly disagree. Means shown.
Overall exam experience by cohort

Typers (left) and Hand writers (right)

Agree (better)

Disagree (not good)

Likert scale: 5 = strongly agree, 1 = strongly disagree.
Running out of time by cohort

Typers (left) and Hand writers (right)

Likert scale: 5 = strongly agree, 1 = strongly disagree. Means shown.
Was the sound of typing distracting?
In each boxplot Typers (left) and Hand writers (right)

The two cohorts ANIM and BIOL were removed from the analysis because typers and hand writers sat in different rooms.

Those that could hear typing (who selected 5, 4 or 3) were included in the determination of distraction by typing sound.

Cohort exams were held in different venues.
Was the sound of typing distracting (VetSci)?
In each boxplot Typers (left) and Hand writers (right)

VetSci Course:

VETS both the internal and external cohorts used same room, but at different times (4 weeks apart).

VETS internal: warm day, ceiling fans and construction noise.
VETS external: cooler day, no fans, quiet.

Environmental conditions and acoustics play a large role in the degree to which ‘typing noise’ becomes a distracting factor.

Hand-writers were not all quiet either!
Future intention to type
Typers (left) and Hand writers (right)


Two cohorts:
BIOL 10 typers, 75 handwrote
OCTY 3 typers, 24 handwrote

Mean shown for each.

For typers ‘yes’ (n = 13).
For handwriters ‘no’ (n = 99).

Likert Scale: 5 = Strongly Agree, 1 = Strongly Disagree
Are some students over estimating the neatness of their hand writing?!

Discomfort from using a pen increased with exam duration (below).

Sig at >.01!

* Note 20% response rate by VETS for this item. All others near 90%
**Trial Technical Issues**

**Issue log:** 15 of the 69 who typed reported ‘technical issues’ via the post-exam survey. 1 more was identified by observation. The majority were minor.

<table>
<thead>
<tr>
<th>Issue</th>
<th>N</th>
<th>Notes, Additional Observations, Suggested Solutions</th>
</tr>
</thead>
</table>
| Boot/start up                | 2  | In reality most participants needed assistance/forgot boot key.  
*Familiarity: need to practice!* |
| Entering ID                  | 0  | All good. (some students entered ‘s’ rather than 8 digit number but system copes fine).                                   |
| Using the software           | 1  | Some did not know how to 'exit' gracefully (i.e. File save, file exit, shutdown).  
*Need to practice! Investigate an 'I’m finished' script/button.* |
| Battery                      | 0  | Most plugged in. *Power needs to be available.*                                                                         |
| Saving files                 | 0  | All good. (noticed one student used ‘save as’ when save was ‘greyed out’) now fixed                                     |
| Software crashed/            | 4  | 1 x Old 2009 white Macbook. Office suite quit to desktop.  
3 x System drive ran out of space causing the system to crash (now fixed). |
| computer froze               |    |                                                                                                                      |
| Touchpad/mouse               | 7  | Sensitivity reported by participants. *Some adjustments were made.*  
*USB wired mice highly recommended!* *Investigate drivers.* |
| Scrolling                    | 15 | Two finger scrolling opposite to OSX, keyboard shortcuts. Small scroll bars. Sensitivity.  
*Familiarity: need to practice. Larger scroll bars. Investigate a user selectable option for touchpad/scroll behavior (and re-mapping of keyboard shortcuts).* |

**Further development** is needed to address these issues.  
Warnings remain in readme files available on public download sites.
Student consideration of general exam conditions when using computer versus pen:
All six cohorts. Response pairs: those who typed (line 1) & those that hand-wrote* (line 2)

* Note - Many of those that hand-wrote their exam had no prior experience of using a computer for an exam so the results presented here are largely speculative on their part.
However, it is reasonable to assume that they drew on their general use of computers.

Note! Updated March 2015 edition places ‘same equally’ in the middle rather than on the right.
Writing strategies under non-exam conditions – general writing habits:
All six cohorts. Response pairs: Typers (line 1) and Hand writers (line 2)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Typers</th>
<th>Hand writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think carefully before I start writing when using my computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I think carefully before I start writing when using pen and paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I take notes in lectures using my computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I take notes in lectures using pen &amp; paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make quick, rough notes before writing essays/reports properly using my computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make quick, rough notes before writing essays/reports properly using pen and paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make a detailed plan before writing essays/reports properly using my computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make a detailed plan before writing essays/reports properly using pen and paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I just start writing (there is no plan!) when using my computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I just start writing (there is no plan!) when using pen and paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make lots of notes using pen &amp; paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tend to go back and re-read and revise my writing quite a lot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I prepare most of my assignments using a computer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nonparametric U & Z used to compare those who typed in the exam to those that hand wrote.

*Note! The September 2014 edition of this chart was incorrectly reversed against the stats – this one is fine!*
Student writing

**Typers (left) and Hand writers (right)**

Likert scale: 1 = strongly disagree, 5 = strongly agree.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Typers Mean</th>
<th>Handwriters Mean</th>
<th>Mann-Whitney U</th>
<th>Z</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I type faster than I handwrite</td>
<td>4.52</td>
<td>3.67</td>
<td>8213</td>
<td>-4.637</td>
<td>&gt;.001</td>
</tr>
<tr>
<td>I type accurately</td>
<td>4.23</td>
<td>3.49</td>
<td>7551.5</td>
<td>-5.089</td>
<td>&gt;.001</td>
</tr>
<tr>
<td>When I make errors, I am able to quickly correct them as part of typing</td>
<td>4.49</td>
<td>3.88</td>
<td>8523</td>
<td>-4.248</td>
<td>&gt;.001</td>
</tr>
<tr>
<td>I often spell detect</td>
<td>3.37</td>
<td>3.61</td>
<td>11097</td>
<td>-1.342</td>
<td>0.179</td>
</tr>
<tr>
<td>I work more efficiently when I type on a familiar keyboard</td>
<td>4.46</td>
<td>4.31</td>
<td>10917.5</td>
<td>-1.656</td>
<td>0.098</td>
</tr>
<tr>
<td>My handwriting is normally neat and legible</td>
<td>3.28</td>
<td>3.48</td>
<td>11621.5</td>
<td>-0.77</td>
<td>0.441</td>
</tr>
</tbody>
</table>
Did the nature of prior experience of e-exams impact on the decision to type this exam?

All participants, all cohorts.

Of those with Prior exp. | All
---|---
Mann-Whitney U | 502
Z | -2.734
Sig. (2-tailed) | >.01

Looks like a ‘yes’!
Does the nature of prior experience of e-exams impact future intended use?

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Hand writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>22.5</td>
<td>21</td>
</tr>
<tr>
<td>Z</td>
<td>-3.262</td>
<td>-2.248</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>&gt;.01</td>
<td>&gt;.05</td>
</tr>
</tbody>
</table>

Looks like a ‘yes’!
Learning more!

How can e-Exams help use learn more?

We will be collecting data electronically – leverage it!

What could we find out?

• Student/cohort performance on each exam question, identify misconceptions.
• Student behaviour in exams – strategies used, time answering questions, marks gained over time.
• Provide rapid and relevant feedback (e.g. strengths/weakness by topic area – helps students know where they need to improve.)
• Quality enhancement of exams themselves - item response analysis to identify weak / low discrimination value questions / options.
Contact:
OLT National Project Leader
Mathew Hillier
University of Queensland / Monash University
m.hillier[at]uq.edu.au / mathew.hillier[at]monash.edu