Progress Report on ALTA Grant: Comparative Evaluation of Marking and Feedback Support Systems

John Venable, School of Information Systems
Acknowledgement

This research is supported by a grant from the Australian Council of Deans of Information & Communication Technology (ACDICT) Learning and Teaching Academy (ALTA)

Thanks!
Grant Research Participants

Research group on Technology Supported Education (TSE)

- Ashley Aitken – School of Information Systems
- Doug Atkinson – School of Information Systems*
- Alistair Campbell – Edith Cowan University
- Vanessa Chang – CBS Teaching and Learning*
- Heinz Dreher – School of Information Systems
- Christian Gütl – Graz University of Technology
- Tomayess Issa – School of Information Systems
- Torsten Reiners – School of Information Systems*
- Brian von Konsky – CBS Teaching and Learning*
- Lincoln Wood – School of Information Systems

* Not named on the grant
Overview

- Marking and Feedback Support Systems (MFSS) Products, Benefits, Features
- Research Objectives
- Research Design
- Research Progress so far
- Research Findings so far
- Research Ongoing
- Research Outcomes and Plans
Marking and Feedback Support Systems (MFSS)

A computer-based information system developed to support (but not fully automate) and improve the efficiency and effectiveness of people marking and providing feedback on essays or other open-ended written assignments.
Burrows and Shortis (2011) identified 15 MFSS (not their terminology)

Commercial products
Blackboard 9.1 Gradebook rubric feature, Waypoint, Moodle Workshops, ReMarksPDF and ReMarksXML, Turnitin’s Grademark, iAnnotate, Review

Research Prototypes
ABBA, EPSS (new version named TTF), tsAAM, Semi-Automatic Essay Assessment based on a flexible Rubric, and SAFS
MFSS Benefits (Intended)

- Improved efficiency of marking and feedback
  Reduced marking and feedback time

- Improved effectiveness of marking and feedback
  Improved accuracy and consistency of marks
  Higher quality feedback
    Legible, Meaningful and comprehensible, Specific to student work, Establishes the relationship to the material taught, Helpful & constructive, Related to the educational goals
  Earlier/Timelier feedback

Improved student satisfaction with feedback
Improved student learning through feedback received
More easily and accurately apply marking guides and rubrics
  Select marks using check boxes or moving sliders

More easily provide written comments
  Pre-written comments, Adding new comments to database, Editing pre-written comments to make them specific to the student’s actual work

More easily place comments at appropriate places on assessed student work submissions

More easily calculate total marks

More easily record marks

More easily produce feedback and summaries

More easily transmit feedback to students
Want to evaluate MFSS in terms of …

1. Tool installation effort
2. Assessment/pre-marking set-up times
3. Resources needed/used
4. Marking effort
5. Marking reliability
6. Accuracy
7. Student satisfaction with feedback

… in comparison with other MFSS
Issues in Research Design

- Confounding Variables in Evaluating MFSS
  - The student – differences in perceptions and expectations
  - The unit – differences in level, topic, and size
  - The assessment item – type, length, assessed aspects
  - The marking rubric – length, detail, quality, feedback text
  - The marker – motivation, experience, domain knowledge
  - Learning effects – how to use the tool, marking the assessment item

- Ethical Issues
  - Inappropriate to assess some students solely with one marking and feedback method/MFSS and other students solely with a different marking and feedback method/MFSS
• Combine field experiment (artificial evaluation) with a field study (naturalistic evaluation)
  • Naturalistic – three realities – real users, real system, real problem/task (Sun & Kantor 2006)
  • Real markers, real students, real system (MFSS), but multiple systems and multiple markers
  • Artificial – more than one system used by each marker, more than one set of feedback received by each student
Control for confounding variables – unit, assessment type, rubric, etc. – by running the experiment multiple times for multiple units, assessments, etc.

Control for student and marker by giving students multiple feedback (different MFSS) and using multiple markers to mark each assessment

Control for learning effects by using MFSS in different order between markers

Include control group – comparison with not using an MFSS (i.e., “manual” marking)
Research Design – Ethics

- All students are assessed using all MFSS (including manually/no MFSS)
- All students are assessed by all markers
- No students are disadvantaged by learning effects
## Research Design (for three MFSS)

<table>
<thead>
<tr>
<th></th>
<th>Students 1-10</th>
<th>Students 11-20</th>
<th>Students 21-30</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marker A</strong></td>
<td>No MFSS (manual)</td>
<td>MFSS X</td>
<td>MFSS Y</td>
</tr>
<tr>
<td><strong>Marker B</strong></td>
<td>MFSS X</td>
<td>MFSS Y</td>
<td>No MFSS (manual)</td>
</tr>
<tr>
<td><strong>Marker C</strong></td>
<td>MFSS Y</td>
<td>No MFSS (manual)</td>
<td>MFSS X</td>
</tr>
</tbody>
</table>
Research Design – Data Collection

- Collect data from all markers on opinions
- Collect data from all markers about marking time and marks
  - Using spreadsheet to record start and stop times
- Collect all the returned marks and feedback
- Collect data from students via survey about opinions
- Designed data collection instruments
- Obtained ethics approval and permission to conduct research with students
- Evaluated in three units
  - Business Information Systems 100 – ABBA vs tsAAM
  - Object-Oriented Development 501 – manual vs tsAAM
  - Supply Chain Information Systems 502 – manual vs tsAAM
- Studies were more exploratory and experience gathering – didn’t fully follow the research design
  - Gathered data on set-up times, marking time, marker experiences
  - Did not mark all assessments using both methods, provide multiple feedback to students, or survey students
Research Findings So Far

- Significant time required to install and set up tsAAM, manage users, etc.
- tsAAM needed customisation to provide some needed features
  - Ability to more fine-tune numeric marks, etc.
- Significant, but reasonable time to set-up marking rubrics, but need technical assistance
- Significant time savings in doing marking (33%)
- Several suggestions for improvements, some made to tsAAM
Evaluating in eight units

Most will follow the rigorous research design

- Business Information Systems 100 – ABBA vs tsAAM vs Blackboard (partial evaluation – too many students)
- Business Problem Analysis 300 – TTF vs Blackboard (2 assignments)
- Business Software Tools 200 – Blackboard only
- Operations and Material Management 301 – tba
- Operations Management 502 – manual vs tsAAM
- Purchasing and Procurement 311 – manual vs TTF (poster present!)
- Strategic Supply Chain and Logistics Management 302 – manual vs TTF
- Systems Analysis & Design 251 – TTF vs Blackboard
Outcomes (So Far and Planned)

- Published paper on the research method and design
- Draft papers underway reporting results of 2011 evaluations
- Plan further papers on individual evaluations
- Plan journal paper integrating and summarising the individual evaluations
- Developing application for OLT grant (submit in August) for further evaluations of commercial products – Blackboard, Review, and Turnitin
Questions and Discussion?

Research Methods

- Design Experiments – Education
  Brown, 1992; Collins, 1992

- Design-Based Research – Education
  Dede, 2004; Dede, et al., 2004; Design-Based Research Collective, 2003; Reeves, et al., 2005

- Design Science Research – Information Systems
  Hevner, et al., 2004; March & Smith, 1995; Venable 2006a, 2006b, Peffers et al 2008
Design Science Research Activities

Technology Invention/Design
Enhancement or creation of a method, product, system, practice, or technique

Theory Building
Solution Space and Problem theories
Utility theories, Design theories

Technology Evaluation
Artificial Evaluation
Naturalistic Evaluation

Problem Diagnosis
Understand problem space
Problem causes and consequences

(Venable, 2006a)
Evaluation in Design Science Research

- **Artificial Evaluation – Rigour sense 1 - efficacy**
  - Computer simulations
  - Role playing simulations
  - Field experiments
  - Lab experiments

- **Naturalistic Evaluation – Rigour sense 2 - effectiveness**
  - Case studies
  - Survey studies
  - Field studies
  - Action research