Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

Assessor efficiency and effectiveness using marking and feedback support systems

Lincoln C. Wood
Curtin University
School of Information Systems, Curtin University, Bentley, WA, Australia

ABSTRACT

Pressures on Higher Education institutions include the need to improve marking and feedback support systems for students while carefully managing costs. Many approaches have been suggested, including the improved use of technology to reduce workload and help provide richer feedback to students. In this research two marking and feedback support systems (MFSSs) are evaluated in an educational experiment. Each assignment was assessed by two markers, using two MFSSs (comparing Spreadsheets and the tsAAM software tool); where half the assignments were marked with one MFSS and the other half with the other MFSS. This enables comparisons of outcomes using different MFSSs while not disadvantaging any student. Neither of the MFSSs enabled faster marking and feedback provision; however, the tsAAM tool allowed the markers to provide significantly more feedback to the students in the same amount of time.

KEYWORDS

Marking and Feedback Support System (MFSS), Assessment, Evaluation, Feedback, Productivity.

1. INTRODUCTION

With increasing pressure on Higher Education Institutions to become more efficient while providing a better learning environment, this research examines two marking and feedback support systems (MFSSs) to understand which allows a marker to become faster and/or more efficient. University education “is a complex process that is as much about whom and how we teach as it is about what we teach” (Leask, 2001, p. 114), with effective feedback as a key ‘how’ in teaching methods. This research aims to support feedback improvements as part of the marking process. This is pertinent as the Higher Education sector is characterised by a paradoxical requirement to improve quality while using fewer resources.

Higher Education institutions are simultaneously faced with increasing staff workloads and challenges in providing adequate feedback to students to support learning. Thus, there is continued challenge of balancing efficiency and effectiveness. This paper investigates the use of marking and feedback support systems (MFSSs) and examines how technology can improve students’ feedback while providing efficiency-based advantages. An initial literature review identifies key topics relevant to technology assistance in assessment. These concepts support the evaluation of an experimental design tailored to ensure no student is disadvantaged. The data are analysed and results discussed, before the implications of the results are outlined and future research directions are presented.

2. LITERATURE REVIEW

The Australasian Higher Education sector saw a dramatic rise in student numbers and concomitant increase in class diversity during the 1990s (Martin and Karmel, 2002) which has made it necessary to “get as many students as possible to meet professionally/academically acceptable levels of performance at as high a level as we can” given available resources (Buckridge and Guest, 2007, p. 144). A key focus in the sector has been on the provision of feedback, identified as a key support for the continued improvement of students (Ricketts and Wilks, 2002). Feedback is provided through classroom learning and active learning, discussions

Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

in class, or through traditional assessment processes. Structuring of activities, incorporating elements of Bloom's Taxonomy (Krathwohl, 2002), also provide unit-level opportunities for improved feedback. Where a course-level approach is taken to focus the structuring of learning activities, the authenticity and real-world relevance can be gradually improved (Reiners and Wood, 2013).

Formative feedback is "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (Shute, 2008, p. 154). Greater formative feedback is linked to focus on the process of teaching and delivery of material which has been observed (De Vita and Case, 2003). Frequent provision of chunks of feedback and engagement through gamification can support this learning process (Wood and Reiners, 2012). Formative feedback can be improved using MFSSs.

MFSSs systems vary in the level of support provided, ranging from fully automatic to providing some support. The automated grading of multiple-choice questions (MCQs) is well-established, but MCQs are not well-received in Higher Education; forcing other approaches to be investigated. Advanced forms of MFSSs include automated grading of essays. Other MFSSs may focus on making the marking process faster, providing feedback more quickly and comprehensively, or in ensuring that the required feedback and 'back office administration' are performed more quickly. Improvements in any of these elements have the potential to significantly reduce the workload of staff members involved in assessing student work.

One of the big benefits of computerised approaches to assessment "include rapid formative feedback to students, reduced marking load for staff, and a closer match between the assessment and learning environments" (Ricketts and Wilks, 2002, p. 478). The decrease in the submission-to-feedback lead time is lauded by Dreher (2006). Other options include different assessment types such as MCQs or modified essay questions (MEQs); these methods are easier and faster to mark but it is perceived that they only test surface-level learning (Palmer and Devitt, 2007). Technology assistance in education is crucial as "innovations designed to strengthen the frequent feedback that students receive about their learning yield substantial learning gains" (Black and Wiliam, 1998, p. 7). Technology-based solutions include bots to engage students, particularly in distance education (Gregory et al., 2011).

Self-assessment tools are viewed favourably by students as they believe it helps to consolidate their knowledge (Ibabe and Jauregizar, 2010). This can supplement and reduce the educators' requirements, so that instructor or educator resources are then freed or released to be deployed elsewhere. This enables an instructor to provide formative feedback using technology supported systems, including **automated systems** allowing entire essays to be assessed automatically using technology (Dreher, 2006, Williams, 2006). **Combined, or hybridised, approaches** use both technological and human-assisted feedback. Taras (2003, p. 562) demonstrated that "student self-assessment with integrated tutor feedback is one efficient means of helping students overcome unrealistic expectations and focus on their achievement rather than on the input required to produce their work".

One method of ensuring that adequate assessment and feedback are provided to students involves the use of technology-supported MFSSs. These can be established and directed to either provide greater levels of feedback using the same amount of assessment time/effort, or to provide the same levels of feedback while reducing the workload on each staff member as each assignment may be marked more quickly.

Prior research on MFSSs includes the study by Burrows and Shortis (2011) on system evaluations which emphasise the integration and overall features in support of evaluation of systems for possible adoption by a university. They use an artificial evaluation, relying on mock evaluation tasks with feedback focusing primarily on the markers' evaluations. Data collection involved open comments and Likert-scale ratings. Heinrich et al. (2012) conducted an evaluation of the support provided by e-learning systems for the management and marking of assessments. They focus on the work flow and the tasks involved in marking and suggest that LMSs should be used alongside specialised MFSS applications.

There are multiple methods that may be employed to bring technology into the educational institutions to support student learning. Through investments in technology support, Higher Education institutions aim to provide superior outcomes for students (a focus on effectiveness), while decreasing staff workload (a focus on efficiency). This allows more students to be accommodated with small, or no, increases in staff levels.

Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

3. METHODOLOGY

This research seeks to evaluate two specific MFSSs in terms of their efficiency and effectiveness in an operational sense. The comparison was required to be made in a way that did not impair any students' opportunities, preventing a classical design using experimental and control groups. A full discussion and rationale for decisions made in the design is addressed by Venable et al. (2012). The research used a design science approach and included 'artificial evaluation' of the MFSSs (Venable, 2006), involving the use of the systems, users, and the actual problem (Sun and Kantor, 2006). This study sought to understand the different tools when they were used in a way that closely mimics real-world use (Fritz and Cleland, 2003, p. 164); an approach that has been applied in research on teaching and learning; e.g., Dede et al. (2004). Thus, the tools were used to mark real student assignments under normal working conditions by markers, requiring intervention and use of the technology. Two MFSSs were investigated in this research.

Spreadsheet productivity support allows comments to be typed and automatically sums marks, reducing the calculation errors. The feedback sheet includes comments and observations as feedback and providing support for the marks provided but this comments box was sometimes left empty. Comments could be cut and pasted, providing a small speed-based advantage over handwritten notes. Processing of results requires each file to be emailed to the student or individually posted to the Blackboard site for the unit.

tsAAM – The tsAAM (technology supported Assignment Assessment and Moderation) prototype was previously developed as part of a Masters Research Project by Daniel Berger (Berger, 2011). The current research involved the installation and adjustment of the prototype to improving the effectiveness and usefulness of the system to the markers. tsAAM allows a click to select a marking bands and include appropriate written feedback for the marking band; previously written comments can be added with a click; additional, specific comments can be typed in and added with a click. This allows a large volume of written feedback to be provided before assignment feedback is exported one-by-one and returned to the student.

Evaluation of MFSSs involves confounding variables which may systemically affect the evaluation results. These include:

- **The student**. Students may have different cultural background which could influence their perceptions or expectations about the type, or quantity, good feedback that they should receive.
- **The students' drive to succeed.** Students that are driven to succeed may make better use of feedback, be more engaged, and be enthusiastic about the feedback.
- **The marker.** Markers differ in terms of their level of motivation (e.g., full-time vs. sessional); experience and knowledge of the domain; and their familiarity with specific questions or topics.

Traditionally, an experimental treatment (in this case, the use of an MFSS) is applied to one group of subjects, while another group (the control group) receives no treatment, enabling comparisons between the groups to demonstrate the effectiveness of the treatment. However, in educational settings this is not tenable as one of the groups may have an advantage due to the treatment. The treatment may be very successful and have positive outcomes with high-quality and valuable feedback, providing an advantage to the learning activities of students in the treatment group. Alternatively, the feedback provided to the treatment group may prove to be of low quality, proving detrimental to the learning activities, and thereby providing an advantage to the control group. The design explicated by Venable et al. (2012) addresses these issues.

The experiment was conducted in Strategic Supply Chain & Logistics Management 302, a third-year undergraduate unit at Curtin Business School (Curtin University) that focuses on strategic decision making using supply chain tools to enhance firm competitiveness. The assessment item was a group-based business report showing the students' ability to analyse a business situation and present effective solutions. There was a 5,000 word limit and a group size of two to four students.

The first marker (Marker A) set the assignment and was intimately familiar with the content, the challenges, and the key issues that would be addressed in an excellent answer. The second marker (Marker B) was a Ph.D.-qualified research associate with significant research and teaching experience in this area.

Each assignment was marked twice with different MFSSs. The student received a final grade and both sets of feedback, ensuring that no student was disadvantaged as part of the educational experiment. Each marker started using one tool and switched to the next tool halfway through the assignments (Table 1). There were 16 assignment submissions. The tests were conducted between the groups of assignments, comparing the time taken by each marker using each MFSS and the amount of feedback using each MFSS. Markers recorded the mark and time taken to mark each assignment. Later, the feedback was processed to extract the

Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

volume of written comments provided to students that could feasibly be used to support their learning. These data were analysed using t-tests to determine whether marking is faster with tsAAM than with spreadsheets and whether there is more feedback provided using tsAAM than with spreadsheets.

Table 1 Example of how the MFSS is determined based on student and marker combinations

	Marker A	**Marker B**
Assignments 1-9	Spreadsheets	tsAAM
Assignments 10-12 & 14-17	tsAAM	Spreadsheets

The study evaluated the MFSSs on two constructs: efficiency and effectiveness. **Efficiency** is a measure of the amount of time, or resources, invested to get the desired output. In this research efficiency was evaluated using the amount of time invested in the actual assessment activities. This excludes consideration of the investment in time taken to learn how to use a given MFSS, process marks into different university databases, or to setup the MFSS for the assignment. **Effectiveness** is a measure of how well the system meets requirements. Thus, the system could be very efficient, requiring little investment of time by the markers to complete the tasks, while also being ineffective as it may provide inadequate feedback to the student. In this research effectiveness is measured by the volume of feedback provided to students, an artificial evaluation.

4. RESULTS AND DISCUSSIONS

The use of independent t-tests rests on the assumptions of normality and equality of variances. The only evidence against the assumption of normality (using the Kolmogorov-Smirnov and Shapiro-Wilk tests) was Marker B's marking time for the second group, where the data were clustered with little variation. The marker recorded time to the nearest five minutes, leaving little variation in the group. There was no evidence against the assumption of equality of variances, except for the volume of feedback between the MFSSs. The results are provided for tests assuming either inequality or equality of variance (Table 2).

Table 2 Independent samples t-test for differences between MFSSs

			Levene's Test for Equality of Variances		t-test for Equality of Means						
										95% CI of the Difference	
			F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Diff.	Lower	Upper
Marker A	Mark	Equal variances (EV) assumed	.303	.591	.801	14	.437	12.73	15.90	-21.37	46.83
	Time	EV assumed	.383	.546	.007	14	.994	0.02	2.17	-4.64	4.67
	Vol.	EV **not** assumed	NA	NA	-5.745	6.248	.001	-715.90	124.62	-1017.91	-413.90
Marker B	Mark	EV assumed	1.559	.232	.701	14	.495	9.16	13.07	-18.86	37.18
	Time	EV assumed	.484	.498	-.963	14	.352	-2.30	2.39	-7.43	2.83
	Vol.	EV **not** assumed	NA	NA	3.940	10.158	.003	217.02	55.08	94.56	339.48

Mark – It was found that there was no significant difference between MFSSs in the number of marks allocated to students (for Marker A: $t(14) = 0.801$, $P = 0.437$; for Marker B: $t(14) = 0.701$, $P = 0.495$).

Time – It was found that there was no significant difference between MFSSs in the amount of time required to assess and provide feedback (for Marker A: $t(14) = 0.007$, $P = 0.994$; for Marker B: $t(14) = -0.963$, $P = 0.352$).

Volume – Using the results of the t-test where equal variances are not assumed, it was found that there was a significant difference between MFSSs in the volume of feedback provided to students (for Marker A: $t(6.248) = -5.745$, $P = 0.001$; for Marker B: $t(10.158) = 3.940$, $P = 0.003$).

The results reveal that when using different MFSSs there was no statistically significant difference in the number of marks when using different MFSSs (this is desired; the MFSS didn't influence different marks to be allocated), nor was there any statistically significant difference in the amount of time taken to assess each assignment. However, we can say with 95% confidence that when using the tsAAM MFSS:

- Marker A generally provided between 414 and 1018 additional words of feedback for each assignment when using tsAAM rather than using Spreadsheets.

Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

- Marker B generally provided between 95 and 339 additional words of feedback for each assignment when using tsAAM rather than using Spreadsheets.

The volume of feedback (an indicator for the effectiveness of the MFSS in providing feedback to students) is significantly different when using different MFSSs. Importantly, neither marker was spending substantially less time per assignment with tsAAM than with Spreadsheets; rather, roughly the same amount of time was invested per assignment, while a greater volume of feedback was provided when using tsAAM.

These results indicate that simply adopting a more advanced MFSS may not, in itself, improve the efficiency of the marking process. It is found that neither tool is faster than the other; rather, tsAAM allows more feedback to be provided in the same amount of time than the use of Spreadsheets does. This conflicts with my initial expectations, that the tsAAM tool would allow faster assessment of assignments.

Neither of the MFSSs impact the time required to read and comprehend each assignment (Marker A has faster comprehension and reading speed; English is his first language). As time is still required to read and comprehend assignments, the overall efficiency gains may be small as the MFSS supports only one task. However, the additional feedback provides a compelling reason for adopting a MFSS. Care must be to ensure processes are designed to enable efficient administration of assignments, marks and feedback in a LMS.

The research has implications for educational managers. While the adoption of MFSSs may be undertaken to reducing staff workload, this research indicates that there may be no reduction in workload unless there is processes ensure the same volume of feedback is provided as with the existing approach, but the work is done more quickly. The improvement in staff workloads may be relatively insignificant and less than anticipated. The adoption of MFSS can be undertaken rapidly in conjunction with existing LMSs to provide greater feedback to students.

The relationship between efficiency and effectiveness in MFSS remains difficult to determine and requires further research to determine how these elements interact. It may even be possible to provide greater volumes of feedback which we may perceive to be ‘shallow’ and yet which students’ perceive to be ‘personalised’, and have students appreciate this less-meaningful feedback more than they may appreciate meaningful, but brief, feedback.

5. CONCLUSIONS

This research investigated the artificial evaluation of two MFSSs used in actually marking assignments, in a way that prevented any student from being disadvantaged during the process. The research shows that MFSSs can increase the volume and value of feedback to students; however, the experiment revealed no significant reduction in the time taken to assess and provide feedback depending on the MFSS used. The initial indication is, however, that the use of MFSS will not, by itself, significantly improve the efficiency of the assessment process, unless *less-valuable feedback* is intentionally provided; i.e., unless the same volume of feedback is provided, in which case this may be completed slightly more quickly. Thus, an efficiency gain may be mandated while using MFSS, which may result in the same value/volume of feedback, but provided in a slightly reduced time. Greater integration of MFSSs into existing LMSs, the use of which is mandated by the university, may provide additional efficiency advantages. While in the past challenges may have been addressed through changing the type of assessment, including MCQs, the future should see further adoption of MFSSs to support assessment. This will improve students’ outcomes while reducing staff workloads, ensuring targeted workload levels can continue to be met while financial constraints are considered.

Limitations include the measure for effectiveness (the volume of feedback); this fails to capture the students’ perspectives on the feedback’s ‘value’. This limitation is addressed with a current survey to gain students’ subjective perception of the feedback (results not available at time of writing). The survey was emailed to students (with a copy of their specific feedback attached to stimulate their memory about the feedback) allowing a naturalistic evaluation to occur. This will improve the measure of effectiveness.

Future research includes the analysis of policies to monitor and control the volume of feedback to manage workloads. This can be accomplished using a system dynamics model. This model may enable managers to establish appropriate policies and understand the impact of different investment decisions.

ACKNOWLEDGEMENTS

Published as: Wood, L. C. (2012). Assessor efficiency and effectiveness using marking and feedback support systems. In P. Kommers, T. Issa, & P. Isaías (Eds.), IADIS International Conference on Internet Technologies & Society 2012 (pp. 237–242). Presented at the International Conference on Internet Technologies & Society 2012, Perth, Australia: IADIS Press.

I acknowledge with thanks the financial support provided by the Australian Council of Deans of Information and Communications Technology (ACDICT); Dr Nasir Uddin (co-marking the assignments); Josef Mosser (technical assistance); and the 'tsAAM research group', which developed the research design discussed in Venable et al. (2012).

REFERENCES

Berger, D. 2011. *Supporting tool for moderation in the grading process of summative assessments: Design and prototype of a software tool for moderation and assessment with variable rubrics.* Masters Thesis, Graz University of Technology.

Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Education*, Vol. 5, No. 1, pp. 7-74.

Buckridge, M. & Guest, R. 2007. A conversation about pedagogical responses to increased diversity in university classrooms. *Higher Education Research & Development*, Vol. 26, No. 2, pp. 133 - 146.

Burrows, S. & Shortis, M. 2011. An evaluation of semi-automated, collaborative marking and feedback systems: Academic staff perspectives. *Australasian Journal of Educational Technology,* Vol. 27, No. 7, pp. 1135-1154.

De Vita, G. & Case, P. 2003. Rethinking the internationalisation agenda in UK higher education. *Journal of Further and Higher Education,* Vol. 27, No. 4, pp. 383 - 398.

Dede, C., Nelson, B., Ketelhut, D. J. C., Jody & Bowman, C. 2004. Design-based research strategies for studying situated learning in a multi-user virtual environment. *In:* Kafai, Y. B., Sandoval, W. A. & Enyedy, N. (eds.) *Sixth International Conference on the Learning Sciences.* Mahweh, N.J: Lawrence Erlbaum Associates.

Dreher, H. 2006. Interactive on-line formative evaluation of student assignments. *Journal of Issues in Informing Science and Information Technology,* Vol. 3, No. pp. 189-197.

Fritz, J. M. & Cleland, J. A. 2003. Effectiveness versus efficacy: More than a debate over language. *Journal of Orthopaedic & Sports Physical Therapy,* Vol. 33, No. 4, pp. 163-165.

Gregory, S., Dalgarno, B., Campbell, M., Reiners, T., Knox, V. & Masters, Y. Year. Changing directions through VirtualPREX: Engaging pre-service teachers in virtual professional experience. *In:* Changing Demands, Changing Directions Proceedings ascilite Hobart 2011, 2011. pp. 491-501.

Heinrich, E., Milne, J. & Granshaw, B. 2012. Pathways for improving support for the electronic management and marking of assignments. *Australasian Journal of Educational Technology,* Vol. 28, No. 2, pp. 279-294.

Ibabe, I. & Jauregizar, J. 2010. Online self-assessment with feedback and metacognitive knowledge. *Higher Education,* Vol. 59, No. 2, pp. 243-258.

Krathwohl, D. R. 2002. A revision of Bloom's taxonomy: An overview. *Theory Into Practice,* Vol. 41, No. 4, pp. 212.

Leask, B. 2001. Bridging the Gap: Internationalizing University Curricula. *Journal of Studies in International Education,* Vol. 5, No. 2, pp. 100-115.

Martin, Y. M. & Karmel, T. 2002. Expansion in higher education during the 1990s: Effects on access and student quality. Canberra: Department of Education, Science and Training.

Palmer, E. & Devitt, P. 2007. Assessment of higher order cognitive skills in undergraduate education: Modified essay or multiple choice questions? Research paper. *BMC Medical Education*, Vol. 7, No. 1, pp. 49-55.

Reiners, T. & Wood, L. C. 2013. Immersive Virtual Environments to facilitate authentic education in Logistics and Supply Chain Management. *In:* Kats, Y. (ed.) *Learning Management Systems: Metrics, Standards, and Applications.* Hershey, PA: IGI Global.

Ricketts, C. & Wilks, S. J. 2002. Improving student performance through computer-based assessment: Insights from recent research. *Assessment & Evaluation in Higher Education,* Vol. 27, No. 5, pp. 475-479.

Shute, V. J. 2008. Focus on formative feedback. *Review of Educational Research,* Vol. 78, No. 1, pp. 153-189.

Sun, Y. & Kantor, P. B. 2006. Cross-Evaluation: A new model for information system evaluation. *Journal of the American Society for Information Science & Technology,* Vol. 57, No. 5, pp. 614-628.

Taras, M. 2003. To feedback or not to feedback in student self-assessment. *Assessment & Evaluation in Higher Education,* Vol. 28, No. 5, pp. 549-565.

Venable, J., Aitken, A., Chang, V., Dreher, H., Issa, T., Von Konsky, B. & Wood, L. 2012. Developing a research design for comparative evaluation of marking and feedback support systems. *In:* Atkinson, R. & Mcbeath, C. (eds.) *Proceedings of the teaching and learning forum 2012.* Murdoch University, Perth, Australia.

Venable, J. R. 2006. The role of theory and theorizing in Design Science research. *In:* Chatterjee, S. & Hevner, A. (eds.) *First International Conference on Design Science Research in Information Systems and Technology.* Claremont, CA: Claremont Graduate University.

Williams, R. 2006. The power of normalised word vectors for automatically grading essays. *Issues in Informing Science and Information Technology,* Vol. 3, No. pp. 721-729.

Wood, L. C. & Reiners, T. 2012. Gamification in logistics and supply chain education: Extending active learning. *IADIS 2012 International Conference on Internet Technologies and Society.* Perth, Australia: IADIS.