



Literature review

on the use of e-learning tools for formative essay-type assessment

Literature review on the use of e-learning tools for formative essay-type assessment

Eva Heinrich¹, John Milne¹, Terry Crooks², Bruce Granshaw³
and Maurice Moore⁴

¹Massey University, ²Otago University, ³Victoria University of Wellington, ⁴Appreciative
Development Ltd

Acknowledgements

Thanks to Professor Luanna Meyer from Victoria University Wellington for feedback on this report.

This project was a collaboration between: Massey University, University of Otago, Victoria University of Wellington and Universal College of Learning (UCOL).

The Tertiary Education Commission funded this project by the e-Learning Collaborative Development Fund.

Correspondence address: Massey University, Institute of Information Science and Technology, Private Bag 11 222, PALMERSTON NORTH

Table of Contents

1	Introduction.....	3
1.1	Literature Review Aims and Methods	3
1.2	Analysis of Literature	4
2.	Essay-Type Assessment.....	6
2.1	Assessment in Relationship to Teaching and Learning	6
2.2	Formative Assessment	7
2.2.1	Relationships between Formative and Summative Assessment	7
2.2.2	The Importance of Formative Assessment	10
2.2.3	What is Needed for Effective Formative Assessment	11
2.3	Assessment with Supply Items	13
2.4	Assessment at NZ Tertiary Institutions.....	15
2.5	Definition of Essay-Type Assessment.....	17
2.6	The Marking of Essay-Type Assessment	18
3	How Computers Can Assist with Essay-Type Assessment	21
3.1	Support Tutors Need in Marking Essay-Type Assessment	21
3.2	Automated Marking Systems.....	23
3.3	Tools to Assist Peer- and Self-Assessment and Detect Plagiarism	25
3.4	Summary	26
4	Tools to Support Essay-Type Assessment.....	28
4.1	Support for Essay-Type Assessment using Learning Management Systems	28
4.2	Support for Essay-Type Assessment Using Generic Software.....	29
4.3	Support for Essay-Type Assessment using Specialist Software.....	30
4.4	Tools Provided as Part of Research Projects	33
5	Issues Concerning the Stakeholders Involved in Assessment	44
5.1	Students.....	44
5.2	Teaching Staff.....	46
5.3	Institutions.....	50
6	Importance of E-Learning Technology for Assessment	53
7	Summary and Preliminary Conclusions.....	55
	References.....	58
	Appendix A – Sources for Literature Review.....	66

1 Introduction

Student's written responses are the basis of much of the learning activity in tertiary education. These include essays, reports, and short answers. The approach that tertiary teachers use to set, mark and manage these types of learning activities has remained relatively unchanged in recent years with the use of e-learning tools making little impact. Students may be expected to submit word processed work and teachers may use an electronic class list to store any marks but little else has changed for most. The focus of this literature review is on the marking and management of student supplied input such as essays, short answers, graphs and computer programmes, something that will be called "essay-type assessment". A comprehensive search of academic journals and conference proceedings was undertaken to establish the status of e-learning use for essay-type assessment in tertiary education.

1.1 Literature Review Aims and Methods

There are two main aims to this literature review. The first aim is to establish the characteristics and importance of formative, coursework assessment. The second aim is to identify e-learning techniques, tools and approaches for this type of assessment, to discuss what is known about their effectiveness and to uncover factors influencing uptake.

For the first aim key articles and books that have shaped current theory on formative assessment of student work have been identified and analysed. For the second aim a comprehensive search of e-learning literature from 2001 to the present was conducted.

A list of about 100 journals and conferences relevant to assessment and e-learning were identified. From these a subset of close to 40 sources was selected, based on journal impact factor, reputation and relevance to the study. Within these sources over 580 journals editions and conference proceedings were searched. The sources mostly reported on work from Australia, USA, UK, Canada and New Zealand. Books on e-learning assessment that were available in the comprehensive education section of the Massey University library were included in the review.

The main method to identify relevant articles was reading titles and abstracts of selected journals articles and papers in conference proceedings. While more time

consuming this method was chosen in preference to keyword search as it presents a more thorough approach to identifying relevant articles. Additionally, keyword search was carried out on several databases. This was done to cover some additional conference publications and to cross-check against the primary search method. Among the keywords used were: online marking, electronic submission, annotation, marking, assessment mark-up, assessment marking, e-learning formative essay, essay mark-up, formative assessment essay, and marking essay online. Appendix A provides information on journals, conferences, databases and books that were considered and used for the literature review.

1.2 Analysis of Literature

At the outset of the review process topic areas, divided into primary and secondary groupings, were defined and consequently refined. These areas were used to categorise interesting articles in preparation of the analysis. The topics of the primary group were:

- Essay marking using some aspect of technology;
- Assignment management tools including learning management systems;
- Assignment marking tools including learning management systems;
- Good practice in assessment and online assessment;
- Institutional acceptance of essay assignment;
- Plagiarism.

The topics of the secondary group were:

- Peer or self assessment using some computing tools;
- Computer mediated communication such as online discussion boards and relationship to assessment;
- Automated marking and management such as computer-based multi-choice questions;
- Automated marking where free-text is marked by the computer;
- Comparison of assessment using different media, e.g. paper approaches, compared to those using a computer;
- Discipline specific marking e.g. computing programmes;
- School context marking and management of essay-type assessment;
- Institutional lessons, such as how to introduce innovations to an organisation;

- Assessment theory, including statements that support assessment in the sense of this project in general and call for systems in support of assessment;
- Learning material or processes derived from assessment.

A relatively low number of articles that could be associated with topics of primary interest were found. This experience matches the opinions expressed explicitly by some authors of articles stating that the higher education sector has not put enough emphasis on the opportunities arising from e-learning and assessment for student learning. Among the articles found for e-learning and assessment the vast majority dealt with assessment via multiples choice tests and quizzes, some articles with the automated marking of essays a limited number with e-learning supported human marking of essays.

After the initial categorisations the identified literature was studied again. An outline for the literature review was defined, articles linked with the relevant sections and the review written.

2. Essay-Type Assessment

This section provides an overview about the characteristics and importance of formative assessment of course work. It starts by setting assessment in relationship to teaching and learning and then moves to formative assessment in general. From there the focus shifts to formative assessment with supply items that are contrasted with restricted response items. After a brief examination of the assessment policies and guidelines of NZ tertiary institutions the term “essay-type assessment” is defined. The section concludes with directions given in the educational literature on the marking of essay-type assessment.

2.1 Assessment in Relationship to Teaching and Learning

Assessment is an important component of education that has to be seen in a wider context of educational goals, course design and student motivation. While a detailed exploration of these issues is beyond the scope of this literature review some high-level introductory comments can be made to provide the context for the following sections.

Education in general and specifically higher education should prepare students for life-long and autonomous learning. Assessment plays a central role in higher education and should facilitate this goal. Assessment must be designed in context of the teaching and learning strategy of a course (Kendle & Northcote, 2000; Macdonald, 2003; Muirhead, 2002). There must be a commitment to a pedagogical rationale for adoption of assessment design (O’Reilly, 2005). Assessment should be student-centred to enable meaningful and relevant learning experiences (Muirhead, 2002). Meaningful and constructive assessments need to challenge students to think critically and should encourage students’ interest in learning (Leathwood, 2005).

As it is widely acknowledged that assessment drives student learning and directs student effort, assessment design must be planned accordingly and must be an integral part of course design (Kendle & Northcote, 2000). Assessment tasks influence the direction and quality of student learning (Maclellan, 2004). To move forward students need to be given more responsibility for assessment processes (Nicol & Macfarlane-Dick, 2006) and must be encouraged to participate (Taras, 2002).

Unfortunately, assessment in higher education does not yet give enough consideration to educational goals, assessment design and impact on students (Nicol & Macfarlane-Dick, 2006; Taras, 2002).

2.2 Formative Assessment

Assessment of student work serves quite a lot of different purposes in the learning and teaching activities of tertiary institutions. These purposes can be classified under two main groupings: formative purposes and summative purposes. The formative purposes involve structuring, guiding, and enhancing student learning. The summative purposes involve certifying and reporting student achievement and admitting students to subsequent learning opportunities.

The prime issue to be addressed in planning, implementing and reviewing assessment procedures is their validity: the extent to which they fulfil their intended purpose. Validity is an overarching issue and encompasses narrower issues such as the reliability and fairness of the procedures.

There are different validity requirements for the formative and summative purposes of assessment. Validity is high for formative purposes when the assessment procedures help the student considerably towards achieving the real goals of teaching and learning. Validity is high for summative purposes when the assessment gives an accurate account of the student's capabilities at the time the assessment occurs, a final grade is awarded, or a selection decision is made.

This project is looking specifically at the use of formative assessment in tertiary education, and how it can be enhanced through the use of e-learning techniques. The goal is to maximize the learning benefit from assessments. That will be the main focus of this literature review. Before concentrating entirely on the literature of formative assessment, however, it is appropriate to consider how formative assessment and summative assessment may relate to each other in a tertiary course.

2.2.1 Relationships between Formative and Summative Assessment

Assessments at the end of a course, such as a final examination, may be seen as largely or entirely summative. This is not entirely true. Even if the students receive no comments from the teacher, and perhaps even do not get their marked work back, the

final examination has a formative effect. Students' motivation, study habits and specific learning activities are affected by the existence, expected nature and content of the impending examination, while the experience of sitting the examination and the result obtained can have a major impact on student perceptions of their ability or self-efficacy in the subject, and on their future study options and choices (Crooks, 1988).

Assessments during the course may have a dual purpose: to count toward the final grade, but also to be used as a source of guidance, direction or re-direction of the students' learning. These usually have all of the formative effects mentioned for final examinations. In addition, however, the students usually get their work back accompanied by comments designed to indicate strengths and weaknesses, and there may be feedback to the whole class based on issues common to several or many students. These are more explicitly intended to be formative.

Another category is assessments during a course that are intended solely for formative purposes, and do not count at all in the final grade for the course. These are now comparatively rare, given the heavy workload of both staff and students in most institutions, but have the clear potential advantage that students know that these assessments are non-threatening and designed to be helpful.

There is an inherent tension between formative and summative assessment purposes, especially if the same assessments are to be used for both purposes (Crooks, 2004). For instance, assessment relatively early in a learning experience may be particularly valuable for formative (learning) purposes but distinctly inappropriate to be counted later for end-of-course grading (summative) purposes. If the early assessment makes a substantial difference to the student's learning, perhaps as a consequence of the feedback resulting from it, that learning gain arising from the assessment means that the summative information from the assessment has become out of date and inaccurate by the end of the course, and therefore inappropriate to be counted in the final grade. The early assessment may have given an accurate picture of the student's performance level at the time of the assessment, but that picture no longer is relevant.

Another aspect of the tension between the formative and summative use of the same work involves the nature of the feedback on the work that the students receive. The feedback may focus on justifying marks or grades awarded, on helping the students do better in the future, or both. Feedback focused on justifying marks strengthens the

summative focus, and is less useful for improving students' work than feedback that makes suggestions about how the work could have been improved. For instance, a comment that paragraphs are poorly structured helps a student less than some examples of how the paragraphs could have been structured more effectively.

A third aspect of the tension arises from preliminary evidence that the mere presence of summative information on student work (such as a grade or a mark) diverts student attention away from the more detailed comments provided. There are numerous reports, from teachers and students, that students often pay little attention to specific feedback if a mark or grade is also provided. These reports are accompanied by a few tantalising pieces of research. Black and Wiliam (1998a) cite the research of Butler (Butler, 1988), who found little learning benefit from feedback that consisted of marks alone or marks together with written comments, but substantial learning benefit where the feedback consisted solely of written comments.

Certain approaches to combining assessment of learning and assessment for learning in the same tasks seem promising. Crooks (2004, pp. 6-7) summarised these:

If students are allowed multiple chances to perform well in a particular area, for instance by being reassessed if they do not meet the desired standard initially, they are very likely to pay careful attention to specific feedback resulting from their initial unsuccessful attempts. Perhaps less strongly, if they are expected to perform several tasks that are similar in nature and demands, the feedback on the earlier tasks is more likely to be attended to than similar feedback on the first of quite unrelated tasks. Another option promoting attention to feedback is subdivision of a relatively large task into a series of smaller stages, with feedback available at each stage but the grade based principally or entirely on the final product. This often involves a tension for the teacher in deciding how detailed and extensive feedback can be before it raises doubt about whether the final product is more the work of the teacher or the student – rather akin to a tightrope walking act, with dangers in both directions. In these examples of approaches that involve both feedback and assessment for credit, I think students can see a close association between attending to and trying to respond to the early feedback and the possibility of greater success on later or more important tasks.

2.2.2 The Importance of Formative Assessment

A number of research and policy papers over the past 20 years have indicated the important impact that assessment practices can have on student learning and motivation (Assessment Reform Group, 1999, , 2002; Black & Wiliam, 1998a, , 1998b; Crooks, 1988; Harlen & Crick, 2003; Hattie, 1999; Madaus, 1988; Shepard, 2000). For instance, assessment can have the following affects on students:

- Focus their attention on important aspects of the subject;
- Communicate and reinforce (or undermine) teaching goals, including key performance criteria and desired standards of performance;
- Give them opportunities to practice skills and consolidate learning;
- Influence their motivation as learners and their perceptions of their capabilities (self-efficacy);
- Guide their choice and development of learning strategies, skills and study patterns.

The substantial effects identified have suggested that it is important to plan assessment carefully so it can be used most effectively to support and improve students' learning and motivation – in other words, used well for formative purposes.

A very influential paper that focused attention on formative assessment is the literature review by Black and Wiliam (1998a), summarized in Black and Wiliam (1998b). Their main conclusion from reviewing ten years of wide-ranging research was that providing high quality feedback on student work is a very powerful way of raising the standard of student work. They stated that "We know of no other way of raising standards for which such a prima facie case can be made." New Zealand professor John Hattie conducted an extensive synthesis of a wide range of educational research, which he summarized in his inaugural professorial lecture at the University of Auckland (Hattie, 1999). He concluded this review of "what works" in educational practice with the comment that "the most powerful single moderator that enhances achievement is feedback."

2.2.3 What is Needed for Effective Formative Assessment

Royce Sadler (1989), in the most influential early paper on the requirements for good formative assessment, identified three elements that are crucial to the effectiveness of formative assessment:

- Helping students to recognise clearly the desired goal (understand what is required), and to appreciate what high quality work looks like;
- Providing students with evidence about how well their work matches that goal, and helping them to develop the evaluative skill to compare with some objectivity the work they are producing in relation to the desired goal;
- Explaining ways to close the gap between the goal and their current performance, and helping them to develop the skills required.

While a major strategy for meeting these requirements is teacher feedback to students, as Black and Wiliam, Hattie and others have identified, such feedback can be made less time consuming and more effective if the nature of the work to be done and the criteria for evaluating how well it has been done are well understood by the students before the work is begun. This requires good explanations by the teacher, and preferably time for the students to explore and discuss the criteria. It is particularly helpful if the teacher can make available examples of similar work at different levels of quality, to illustrate the qualities that are being sought (Sadler, 1987). With these conditions in place, the feedback can be much more focused, aimed at fine-tuning the students' understanding of the desired qualities and how to adjust their work to better exemplify those qualities.

Sadler (1989) argued that self-assessment is a vital component in learning. He said that if students were to be able to improve, they must have the capacity to monitor the quality of their own work during its production. Feedback on assessment cannot be effective unless students accept that their work can be improved and identify important aspects of their work that they wish to improve. If students are asked and encouraged to critically examine and comment on their own work, assessment can become more dialogue than monologue, and can contribute powerfully to the educational development of students. As Wynne Harlen and Mary James (1996) put it,

students have to be active in their own learning (teachers cannot learn for them) and unless they come to understand their own strengths and weaknesses, and how they might deal with them, they will not make progress.

Student motivation is crucial to learning. Assessment is one of the major influences on student motivation (Crooks, 1988; Harlen & Crick, 2003). It is important, therefore, to anticipate and try to optimise the motivational effects of feedback on assessment. The research evidence available suggests that the greatest motivational benefits will come from focusing feedback on:

- The qualities of the student's work, and not on comparisons with other students;
- Specific ways in which the student's work could be improved;
- Improvements that the student has made compared to his or her earlier work.

The formative functions of assessment will be seriously undermined if the feedback on assessment does not consistently and strongly emphasise the most important aspects of student learning and development, or if it gives undue emphasis to less important aspects (Black & Wiliam, 1998a; Crooks, 1988). Consequences of such failure are poorly directed student effort, and hence lower validity for formative purposes. For instance, if attention to assignment content is neglected in favour of a strong focus on surface features such as presentation, spelling, punctuation and grammar, the latter may improve but the original purpose of the assignment is not encouraged and supported through appropriate guidance.

Most writers on formative assessment or feedback on student work indicate that a good balance between criticism and positive comments is most effective. Students pay attention to comments that they perceive to be useful and constructive. Comments on strengths make them more receptive to other feedback, while the most effective comments on weaknesses not only identify them but also suggest how they might be addressed.

Nicol and Macfarlane-Dick (2006) emphasise the importance of learner self-regulation. They suggest seven principles of good feedback practice:

1. Helps clarify what good performance is (goals, criteria, standards);
2. Facilitates the development of self-assessment and reflection in learning;
3. Delivers high quality information to students about their learning;

4. Encourages teacher and peer dialogue around learning;
5. Encourages positive motivational beliefs and self esteem;
6. Provides opportunities to close the gap between current and desired performance;
7. Provides information to teachers that can be used to help shape teaching.

Higgins et al. (2002) report that students, despite paying lots of attention to grades, have an intrinsic motivation to learn from feedback. The potential of influencing student learning with formative feedback is large. To achieve this, timely feedback is vital, must explain misconceptions and must suggest pathways to improvement.

2.3 Assessment with Supply Items

The previous section has looked at the purpose of assessment, discussing the characteristics of and differences between formative and summative assessment. The importance of formative assessment, as well called 'assessment for learning', has been emphasized. This section shifts the focus from purpose of assessment to forms of assessment and investigates the learning outcomes targeted by the respective forms.

Many different forms of assessment are available. (Lambert & Lines, 2000), for example, describe the main categories of assessment as classroom observations, objective tests, essays, structured questions, enquiries, self-assessment, and oral assessment. Underlying these categories are a variety of overlapping ways of classifying assessment. In classroom observation assessment is a by-product to a learning activity whereas in a test assessment is the primary objective. In contrast to peer- or teacher-assessment in self-assessment assessor and assessee are the same person. Oral assessment implies a dialogue between assessor and assessee while in forms like essays there is a fairly strict sequence of separated phases of setting a task, producing an artefact as response to this task, followed by the assessment of the artefact. In most cases the artefact is a written document but it can as well be a recording of a presentation or performance.

Common in the literature is the distinction between 'supply items' and 'selection-type items' (Gronlund, 2006; Hanna & Dettmer, 2004; Linn & Miller, 2005; Nitko, 2004b). Selection-type assessment items imply a selection from predefined choices whereas supply items require students to retrieve material from their own memory and

to compose answers in their own words. Supply-items provide freedom of response. A common format of supply items are essays.

For essays the assessment literature distinguishes two forms of questions, namely restricted response questions and extended response questions (Gronlund, 2006; Hanna & Dettmer, 2004; Linn & Miller, 2005; Nitko, 2004b). The distinction relates to the form and scope of the answers expected and the degree of freedom given to the student. Restricted response questions target more narrowly defined problems and limit the answers in terms of format and lengths. Extended response questions might pose some practical limitations, as the maximum number of words for an essay, but otherwise give the student freedom of form and scope. Restricted response questions aim at measuring comprehension and application whereas extended response questions target synthesis and evaluation. Extended responses are well suited to evaluate understanding of concepts and the ability of expressing these (Committee on the Foundations of Assessment, 2001).

Essay questions provide students with freedom of response and require students to apply their knowledge and skills in organizing, integrating and evaluating information (Linn & Miller, 2005; Nitko, 2004b). Essays target knowledge and understanding and proficiency in reasoning (Stiggins, 2005). Essays aim at the highest level of learning outcomes by emphasising the integration and application of ideas (Gronlund, 2006). In contrast to multiple choice examinations essay-type evaluation prompts students to adopt deep learning approaches (Struyven, Dochy, & Janssens, 2005).

As supply items essays require the students to formulate their own responses. Writing requires thinking and reflection to be able to communicate knowledge in clear, plausible and effective ways (Tynjala, Mason, & Lonka, 2001). Data and ideas need to be organized and integrated (Linn & Miller, 2005). Writing is integral to the process of learning (Goodfellow & Lea, 2005). It plays an important part in language development and knowledge construction (Lindblom-Ylanne & Pihlajamaki, 2003). With these characteristics writing plays an important part in achieving higher level learning outcomes.

Using essays for assessment is of value in all subject areas (Linn & Miller, 2005). In disciplines like the social sciences and humanities writing has always been part of assessment. But also in the sciences and in engineering writing has proven to be

beneficial for the understanding of scientific concepts (Aller et al., 2005; Hand, Prain, & Yore, 2001; Tynjala, Mason, & Lonka, 2001).

While writing and the textual format are probably most common, there are other ways to display data and ideas in the sense of supply items for higher level learning. In disciplines like computer science, engineering or architecture the development of models, expressed via diagrams, forms an essential part of the learning process. Like the writing of essays the construction of a model requires analysis, integration and evaluation. Depending on discipline and task specification, an essay can contain combinations of written text, calculations, pictures and diagrams.

2.4 Assessment at NZ Tertiary Institutions

In this section publicly available documents on assessment policies and guidelines of New Zealand tertiary institutions are reviewed. These documents show a distinct difference between polytechnics and universities. The polytechnic documents focus on assessment procedures and regulations whereas the university documents contain considerable amount of information on assessment in regard to learning outcomes.

Assessment guidelines and policy documents of ITP New Zealand (2006), Nelson Marlborough Institute of Technology (2005), Tai Poutini Polytechnic (2006), The Open Polytechnic of New Zealand (2006), Universal College Of Learning (2006), Waikato Institute of Technology (2006) and Waiariki Institute of Technology (2006) have been reviewed. These documents all focus on procedural issues around summative assessment. Procedural issues are outlined, for examples the timeframes for informing students about the assessment components for their courses, requirements for attendance and opportunities for re-sits and exemptions. In terms of character of assessment items distinction is made between achievement-based and competency-based assessment items. Internal assessment and externally moderated assessment is defined. Quality of marking is indirectly addressed by requiring staff to make marking criteria available and by outlining procedures for moderation. Feedback and learning from assessment are mentioned only in passing. Brief statements are made requiring feedback on assignments to be constructive and to show how the work could be improved.

The assessment documents available from the New Zealand universities have a very different focus. For most courses in universities assessment consists of two parts, coursework (as well called internal examinations) and final examinations. Final examinations are always of summative nature whereas coursework combines summative and formative aspects. Coursework can take a variety of forms. For example, the University of Canterbury names “essays, reports, exercises, and tests” as possible forms in its Assessment Guidelines (University of Canterbury, 2004, p. 2). In their guidelines or policies all universities cover issues like the need for linking the assessment to learning outcomes, the recommended distribution of coursework to final examinations, and the need for ongoing staff training. Of special importance for this project are the emphases on feedback and on depth and types of learning.

In regard to feedback the universities state:

- “In addition to receiving a mark or grade, students receive a clear indication of the strengths and weaknesses of their work, accompanied by guidance on how to perform the task or a similar task better.” (University of Auckland, 2002, p. 5)
- “Good teaching according to established best practices in higher education includes provision for feedback to students regarding their developmental mastery of the learning outcomes for specific course content. Such feedback enables students to self-regulate learning behaviour and provides students with fair opportunity to adjust learning behaviour towards the expected mastery of learning outcomes.” (Massey University, 2004, p. 2)
- Feedback needs to be “sufficiently detailed, providing clear indications of the strengths and weaknesses of the students' work and clear guidance on how to perform better on future tasks of a similar nature;” and “constructive, helping to motivate the student to put further effort into learning.” (University of Canterbury, 2004, p. 3)

In regard to depth and types of learning the universities demand:

- “Setting assignments and exam questions which encourage independent thinking.” (University of Canterbury, 2004, p. 12)
- “Assessment in all courses should give explicit emphasis to long-term goals, such as the development of written and oral communication skills, research skills and thinking skills in addition to short term goals.” (University of Canterbury, 2004, p. 4)

- “Assessment practices should encourage deep learning as appropriate for vocational and professional practice areas.” (Auckland University of Technology, 2002)
- “Develop assessment practices which reward demonstrations of intellectual independence.” (University of Otago, 2002, p. 5)

Combining the demands for detailed feedback, that provides guidance towards improvement, and for assessment, that encourages deep learning and intellectual independence, points towards specific forms of assessment. Firstly, such assessment has to be part of the coursework as only coursework has formative character. Secondly, among coursework one needs to look for forms of assessment that encourage students to research problems, construct solutions and formulate their own answers. Assignments typically fulfil these criteria. Students are given a problem specification. They explore the problem and develop a solution which they then deliver in form of an essay, report or oral presentation. This form of assessment stands in stark contrast to tests. While part of course work, tests typically are presented in a question and answer format. The student is presented with a large number of questions that usually require fairly short answers. Often, the student is given a selection of possible answers from which to choose the correct ones. While a valuable form of assessment tests do not fulfil the demands specified earlier.

2.5 Definition of Essay-Type Assessment

The previous sections of this review have given a brief insight into the characteristics of assessment and its goals in the context of teaching and learning. It has been shown that a variety of terms are used to distinguish the various forms of assessment. While all these terms and their underlying concepts have merit, none represents the focus of this research one hundred percent. To facilitate common understanding and clear communication this section motivates and defines the term ‘Essay-type assessment’. Essay-type assessment has the following characteristics:

- Course work: This project deals with course work and not final exams. Course work implies that the learner receives feedback. Course work targets important skills such as deep learning and intellectual independence. Common forms of course work are tests and assignments. The term test carries the connotation of restricted response and summative focus, whereas assignments are commonly

linked to open response and a dual formative and summative character. As the latter term matches more closely the direction of this project this term is chosen.

- Formative: This project focuses on formative as compared to summative assessment. While in practice most formative assessment has as well a summative component, the formative aspect is credited with facilitating learning. If done well formative assessment has a strong positive effect on student learning and motivation.
- Supply-item: This project works with supply-items and more specifically with open-response items like essays. These assessment items require the students to construct and formulate their own responses. They facilitate the development of important skills such as analysis, integration and evaluation. The contrasting assessment form of selection-type items does not provide these learning opportunities.
- Artefact: This project concentrates on assessment that results in the student producing an artefact for assessment. Such an artefact can be a traditional essay, but can as well consist of or contain diagrams, calculations or drawings. The distinction between 'essay' and 'essay-type' is made to include assessment in a wide range of disciplines, such as the sciences, that do not normally use the term essay for their student work. Assessment of presentations, performances or practical work is not the focus of this project. While such assessment will share many of the characteristics of the assessment of artefacts as outlined here, the practicalities would be quite different.

2.6 The Marking of Essay-Type Assessment

One of the practical challenges around essay-type assessment is the time consuming nature of the marking (Linn & Miller, 2005) that has to be performed by a human expert (Hanna & Dettmer, 2004). The individual nature of essays and the types of learning outcomes targeted make it largely impossible to define one correct answer. Automated scoring of essays and the automated provision of individually targeted formative feedback is on the whole not possible (see Section 3.2 for details).

The nature of essay-type assessment poses a range of challenges for the marking process. Knowledgeable human markers are required for judging the quality of responses (Hanna & Dettmer, 2004). The marking of essays is very time consuming

and the reliability of the marking can be very low (Linn & Miller, 2005; Nitko, 2004b). To address these challenges the literature presents a very coherent picture of suggested techniques and procedures (Gronlund, 2006; Hanna & Dettmer, 2004; Lambert & Lines, 2000; Linn & Miller, 2005; Nitko, 2004a).

Successful marking starts with the definition of the essay questions. These questions need to be clearly linked to learning outcomes. It has to be assured that the most suitable form of assessment is chosen for the learning outcomes targeted. Next, the marking has to be guided by a scoring rubric. There are two main forms of such rubrics that support analytic and holistic marking. An analytic scoring rubric requires an outline of an ideal answer or a list of major elements that should be included in an answer. It specifies the number of points that can be achieved for an answer or part of an answer. Restricted response items are commonly judged with the help of analytic scoring rubrics. Based on the nature of extended response questions and the inherent degree of freedom it is not possible to provide a single model answer. Holistic marking rubrics are suggested to assist in the marking of such questions. This involves the definition of quality criteria by which answers are judged and the provision of scores for each quality level. The availability of a scoring rubric allows the marker to focus on the learning outcomes to be assessed. Without guidance from a rubric the marker can be influenced by matters of presentation over content.

Ideally an appropriate scoring rubric should be developed before administering the essay questions. Planning of the scoring will help to refine the questions. Making holistic scoring rubrics available to students before they write their essays will allow them to focus their efforts into the right direction.

A number of procedures are suggested for the marking process. If an assessment consists of multiple questions each question should be marked separately. This will prevent the so called 'halo' effect, where a marker judges the merit of an answer not only based on its merit but influenced by good answers in the same essay to other questions. Focusing on each question separately implies that the marker has to work through the essays multiply times. If this is done it is best to vary the sequence in which the essays are assessed. It has been shown that the judgement of markers can evolve over the course of marking essays of a whole class. A marker might assess the same essay differently, depending on when the essay is looked at. This is called 'marker drift'. To counteract marker drift an essay that has been marked early on

should be marked again by the same marker later in the process. If possible, the assignments should be marked anonymously, that is the marker should not know the identity of the student who has submitted the assignment. Ideally, an assignment would be marked by more than one knowledgeable marker. Following these procedures and using an appropriate scoring rubric will greatly enhance the reliability of essay marking and will further save time in making.

Returning to the formative aspects of essay-type assessment, the importance of feedback needs to be emphasised. The marker should provide feedback to each student, outlining strengths and weaknesses in their work and guiding towards further learning (Linn & Miller, 2005; Nitko, 2004b; Torrance & Pryor, 1998; Tynjala, Mason, & Lonka, 2001). Individualised feedback that provides detailed information on the quality of an answer is mostly given in conjunction with an analytic scoring rubric. Even when using a holistic rubric individual feedback should be supplied (Nitko, 2004b). While facilitating student learning is the most important aspect of feedback a further advantage are the conclusions that can be drawn for teaching. Collecting all feedback the marker can identify strengths and weaknesses of answers across the whole class. This information can be used as a guide for further teaching (Nitko, 2004b). Assessment of essays and especially the provision of individual feedback are very time consuming. The development of a statement bank of frequently used comments can make this process more efficient (McLachlan-Smith & Irons, 1998).

3 How Computers Can Assist with Essay-Type Assessment

Tutor marking is a necessary part of essay-type assessment. In tertiary education the majority of this marking is done by hand with little help from computers (Denton, 2003). There are tools that assist tutors with the management and marking of essay-type assessments. This section will outline some of these tools and discuss the types of tasks that computers can help tutors with in marking assignments.

There are marking tools that once setup automatically provide feedback to students. The two main groups are the restricted response or computer-assisted assessment (CAA) tools and the automated tools for free-text and other types of assignments. The reasons why these types of fully automated tools are not sufficiently robust for essay-type assessment are discussed.

3.1 Support Tutors Need in Marking Essay-Type Assessment

Stephens, Sargent, & Brew (2001) suggest features tutors would need in an ideal computer marking and management tool. The items suggested are:

- A way of providing feedback that is easy to use e.g. an electronic pen;
- Support of all types of submission including electronic student submission of various file types;
- Centralised pre-testing for plagiarism;
- A comments bank that is easy to edit and develops during marking;
- The support of marking schemes that vary from simple to sophisticated;
- Automatic grade and report computation;
- Electronic feedback via student email;
- Electronic storage of marked work and lecturer's comments, marks and annotations;
- Easy links to university central systems.

Similar models have been suggested for computer science projects (Vastani, Edwards, & Pérez-Quñones, 2005) and learning management systems (Jafari, McGee, & Carmean, 2006). The key messages these authors add were to emphasis compatibility and interoperability with existing systems such as student records, and that the systems should understand the individual preferences of learners and tutors.

Baillie-de Byl (2004) reports on an electronic system to assist teams mark and manage assignments. This system provides electronic submission and return of assignments, management of grades, marker team management, and assistance with the marking process. Similar systems have been reported (Campbell, 2005; Denton, 2003; Edwards, Fernandez, Milionis, & Williamson, 2002; Moreale, Whitelock, Raw, & Watt, 2002; Plimmer & Mason, 2006; Wells, 2006; Zhang & Heinrich, 2005a, , 2005b).

Although these systems are similar there are variations, like in the way the tutor enters comments or feedback. Some systems allow the tutor to provide comments directly in the assignment, similar to the tracking function in Microsoft Word, others allow the tutor to store the comments in a different file or append the comments to the end of the student work and describe the location to which the comment refers. Plimmer & Mason (2006) describe a system that digitally annotates a document with a hand-held pen and a tablet laptop. This is described as digital ink and looks similar to hand-written comments. All these systems have their strengths and weaknesses. The appropriateness of a tool depends on the specific marking situation and personal preferences.

The advantages of such tools in general to assist tutors include: improved legibility (Bridge & Appleyard, 2005); staff can edit the feedback as they work through the assignments (Bridge & Appleyard, 2005); tools help marking teams to be consistent as the markers can see each others comments; markers individually marking an assignment can check feedback and identify if it has changed as the marking proceeded (Barrett & Luca, 2002; Campbell, 2005); the tools can save tutors time in the marking and management of assignments (Baillie-de Byl, 2004; Denton, 2003).

The disadvantages of the tools are: tutors are concerned about spending too much time working on a computer (Bridge & Appleyard, 2005); staff development is required for the new systems (Denton, 2003); and staff may not want to change to a new system even when there are efficiency and effectiveness gains (Denton, 2003).

Specific tools are introduced and discussed in Section 4 of this review.

3.2 Automated Marking Systems

Computer-assisted assessment (CAA) is defined by Sims and colleagues (2006) as “the use of computers to deliver, mark and analyse assignments or exams”. While this definition is wide enough to cover both restricted response and supply items, CAA is most commonly used in connection with restricted response items. For restricted response items, such as multiple choice tests, questions, distracters and correct answers as well as feedback can be pre-defined. These tests can be used for both formative and summative purposes, and, once setup, can be used by students independently of teacher involvement. CAA tools for this type of assessment are readily available, yet barriers still exist regarding the time required to master the tools, the skills needed to write good assessments and a lack of institutional strategy (Sim, Read, & Holifield, 2006). This review will not consider these types of CAA tools any further as it targets supply and not restricted response items.

A number of researchers have developed CAA tools for the automated marking of supply items like short answers and essays (Dikli, 2006; Mason & Grove-Stephensen, 2002; Valenti, Neri, & Cucchiarelli, 2003). While computers are good at judging restricted response questions they perform far less well when marking items supplied by the student. For essays CAA tools are not rigorous enough for assessing language development and knowledge synthesis and evaluation. Christie (2003) describes automated marking of free-text as at a similar stage of development as the first powered flights “not very grand, not too elegant and a bit under-powered”. These systems generally need to be setup and calibrated with a set of marked work and feedback (Dikli, 2006). This means that it is only worthwhile for large classes that repeat the same assessment for a number of cohorts.

Besides written language supply items might include elements like graphs, diagrams or computer programmes. These elements have to be treated differently than text. There are specialised automated tools that provide computer generated feedback on concept maps (Conlon, 2006), spreadsheets (Blayney & Freeman, 2004), computer programming (Ala-Mutka, Uimonen, & Järvinen, 2004; Korhonen, Malmi, Nikander, & Tenhunen, 2003; Kutay & Ho, 2005), online discussions (Wu & Chen, 2005), learning portfolios (Chen & Chen, 2005) and language learning (Liu, Wan, & Chen, 2005).

It has been reported that the acceptance of automated essay scoring systems is growing (Rudner, Garcia, & Welch, 2006). The potential of these systems is to save time and costs in marking, to make marking more reliable and generalisable, and to provide immediate feedback (Dikli, 2006).

In the eyes of some researchers automated marking systems do have a place, if they are used playing to their strengths. Mason & Grove-Stephensen (2002) suggest that students critically assess the feedback they get from the computer and understand that computers do make mistakes. The automated systems could be used to assist staff assess rather than be considered as a tool that provides the correct answer (Dikli, 2006).

Conlon (2006) outlines a tool to analyse concept maps. The student map is compared to the expert map and feedback is generated. It is acknowledged that the system makes mistakes and is not sophisticated enough to fully assess concept maps. This weakness has been turned into a design feature and students are reminded that the system is fallible and that they are to use the feedback as a starting point for investigating how they could improve.

Valenti, Neri, & Cucchiarelli (2003) reviewed ten automatic free-text marking systems that assessed text using natural language processing or statistical techniques. They report relatively good correlation between the systems and human markers, however there are major problems. These include problems analysing sentence structure and the huge variations in using free-text to express the same concepts. The major problem is obtaining a large enough sample of marked assignments on which experts agree that can be used to train the automated system. Some automated systems allow the human marker to supplement or change the automatically generated feedback. Other systems provide the student with immediate feedback that they incorporate into their final draft before submitting to a human marker. This emphasises the need for human markers.

With further development automated marking systems for essays might offer opportunities in the future. At this stage these tools do not provide acceptable solutions for essay-type assessment, as the validity of the marking these tools provide is not sufficiently guaranteed, as they are only applicable in situations with very large numbers of student submissions for the same essay question, and as they do not

address a mixture of textual and non-textual student work. For these reasons these types of CAA tools will not further be considered in this review.

3.3 Tools to Assist Peer- and Self-Assessment and Detect Plagiarism

The educational value of self- and peer-assessment designs has been well established in the literature. A number of researchers have described the use of e-learning tools in this area (Barrett & Luca, 2002; Davies, 2002; Davies, 2003; Freeman & McKenzie, 2002; Ge & Er, 2005; Lindblom-Ylänne & Pihlajamäki, 2003; Liu & Tsai, 2005; Peat & Franklin, 2002; Sung, Chang, Chiou, & Hou, 2005; Wong, Wong, & Yeung, 2001). The approaches developed can be divided into two main groupings. Firstly, there are approaches that are built around a traditional artefact like an essay. Tools are used to manage access to and exchange of artefacts, assign student peer markers, collect, distribute and discuss feedback. The tools required to facilitate these tasks have very similar characteristics to tools used by teachers, tutors and marking teams. What has to be carefully considered is how roles, rights and responsibilities are assigned to staff and students. Self- and peer-assessment approaches around essays are very relevant to this research and suitable tools are covered in Section 4 of this review. An example of a peer-assessment design around essays has been given by Barrett & Luca (2002) who devised a system where student assignments, marks and comments are made available for all students to review and reflect on. Students can also provide comments on their classmates work. A main advantage with this open system is that tutor feedback can become more consistent across tutorial groups. Davies (2003) reports on a system where students provide anonymous comments and where they can have anonymous discussions on the comments. It is important not to lose sight of the fact that the effectiveness of peer- and self- assessment is dependent upon how the systems are used (Lindblom-Ylänne & Pihlajamäki, 2003). This includes using these tools to actively construct knowledge and learning skills.

The second group of approaches to self- and peer-assessment has a very different starting point. These approaches do not centre around single assessment artefacts but on a communication paradigm. Students and tutors engage in online discussion forums or chat sessions. They exchange information and viewpoints, and, in a way, peer-feedback is contained in each of these exchanges. The tools used are mostly

standard components of web-based learning management systems, called discussion boards or forums and online chat. These approaches clearly have potential in terms of student learning. Yet, as they do not relate to essay-type assessment, they are not specifically considered in this review.

Academic dishonesty and plagiarism occur frequently in tertiary education (de Lambert, Ellen, & Taylor, 2006; Parameswaran & Devi, 2006) and it is more common in essays and programming assignments (Alam, 2004). Detection systems along with telling students what academic dishonesty is, and promoting values that institutions want, are recommended ways to manage this issue (de Lambert, Ellen, & Taylor, 2006; Macdonald & Carroll, 2006). Submitting work electronically will allow detection systems to be used. It will also allow students to submit work to the detection system for checking and then act on the feedback the detection system highlights before submitting the work for marking and human feedback.

A review of detection systems identified the submission process as being difficult to use, requiring what was described as a fair degree of technical competence (McKeever, 2006). It is important that students and staff can easily use tools. The need to comply with data protection legislation while using these tools is very real issue (McKeever, 2006).

Approaches and tools to prevent academic dishonesty are certainly relevant in the context of essay assignments. Some specific tools will be reviewed in Section 4.

3.4 Summary

Students are concerned about the limited amount of feedback and that their marks are sometimes unfair (Campton & Young, 2005). Tools to assist human markers have been reported to save tutors time so may allow tutors to spend more time providing feedback. The tools can improve the perception of fairness by allowing for a greater sharing of information among markers and with students. Providing electronic submission of assignments will allow the use of plagiarism tools and opportunities for the management of peer- and self-assessment.

McKenzie (2004) quotes a student who said “Quality is a property of the marker, not the system”. While it is true that the marker still provides the feedback and marks and the quality of their input will vary, the system can assist the marker improve the

quality of their work by removing repetitive jobs and allowing a more transparent marking system.

4 Tools to Support Essay-Type Assessment

This section discusses learning management systems and their capabilities to assist the management of marking. The role of generic applications such as Microsoft Word, PDF documents, and Excel to help lecturers mark is explored. Specialist software is looked at to investigate if and how these tools meet the needs of lecturers. This review of tools includes plagiarism detection software.

4.1 Support for Essay-Type Assessment using Learning Management Systems

Today's widely used learning management systems (LMS) such as WebCT (2006), Blackboard (2006) and Moodle (2006), provide basic support around the management of assignments. These systems allow the lecturer to setup an assignment definition, the students to submit their assignments, the lecturer to access the student submissions and to return marked assignments, and the students to retrieve their marked assignments. Setting up an assignment definition includes the specification of parameters for due dates, extension periods or multiple submissions. The LMS relieves the individual lecturer of the responsibilities for secure access and safe storage. The student prepares the assignment outside the LMS, using general purpose tools, submits the assignment and receives results via a web-browser interface.

While the support given is helpful to lecturers some of the functionality provided by LMS is not easy to use or is limited. For instance, the return of marked assignments to students requires very repetitive work in WebCT, as the assignments have to be returned individually to each student. More advanced tasks, like the allocation of markers to individual assignments or the exchange of marking information among markers, are not handled by these tools. There is no way to specify which students have submitted an assignment together and to return the appropriate feedback to all members of a group. One cannot indicate the marking status of an assignment. Labelling an assignment during marking as 'in progress', 'marked' or 'revisit' would be a good way to facilitate the marking process, especially when multiple markers are involved. Most critically, little support is provided for the actual marking task. In WebCT the assignment tools provide markers with the facility to enter comments for assignments. This is done via a comment field that refers to an assignment in its

entirety. There is no way to link a comment to a specific part of an assignment. Scoring rubrics, that form an essential part of educationally sound marking, are not supported. The assignment tools of Blackboard and Moodle have very similar restrictions.

4.2 Support for Essay-Type Assessment Using Generic Software

Marking of electronic copies of assignments directly and annotating these with feedback is possible with general purpose tools like word processors or PDF writer software.

Word processors can be used to write documents for the essay task description, the scoring rubric and general feedback. They can also be useful for commenting on essays. For example, Microsoft Word provides two methods to add comments to essays. Comments can be placed in one specific screen area, e. g., below the document, and the reference between comment and essay is created by highlighting the reference point in the text. These comments can be in text or even audio format. Another form of commenting is the track changes mechanism. The advantage of track changes is that the comments appear directly beside their reference points. The disadvantage is that this method changes the format of the essay. Disadvantages of both approaches are that they require modifying the file submitted by the student so there is no protection against inadvertently changing some of the student's work. Also, both approaches are tightly linked to the structure of the assignment document. Highlighting areas outside given paragraph structures to make higher level comments is not possible.

Adobe Acrobat Professional (Adobe, 2006) allows comments to go directly into PDF documents. This is good as the comments are placed in separate textboxes that are visible right beside their reference points in the essay, without changing the essay layout. It is not possible to inadvertently modify the student's work. Additionally, the comments can be hidden from the essay view and comment summary pages are available.

A spreadsheet program like Microsoft Excel allows markers to record and track assignment details. Many course administrations use the spreadsheet functions to enter class lists, the marks and other information about the student assignments. With

knowledge of the VBA language it is possible to customise Excel to link it to other systems and automate some tasks.

All these generic approaches provide a base level of functionality and their clever use can support marking to a certain degree. Frequently-used comments can be collected in a separate document and suitable comments can be pasted into an essay. For each student a separate summary sheet can be written, containing a copy of the scoring rubric, an overall comment for the essay and the marks awarded.

The disadvantage of generic tools is that they do not provide any of the specifics demanded in the assessment theories. As a consequence, the individual marker needs to spend considerable effort in customising the generic tools and then in living with the work-arounds while marking each essay. There are two main problems with this approach. Firstly, the markers need to be very familiar with the assessment theories as they do not receive any guidance from their tools. Secondly, marking, especially formative marking, is an inherently time-consuming activity. Without efficient tool support it is likely that markers will restrict their efforts in commenting and interacting with co-markers to a minimum. What suffers is the quality of marking and opportunities for learning and teaching are missed. While working with generic tools can provide valuable support to teachers and markers, the use of specialised marking applications should have preference.

4.3 Support for Essay-Type Assessment using Specialist Software

Dedicated marking applications provide lecturers with more specialised support than LMS or generic software. Specialised tools like Markin (Creative Technology, 2006), GradeMark (Turnitin.com, 2006c), Re:Mark (MyDropBox.com, 2006a), WebCTConnect (Massey University, 2006b) or MarkTool (Massey University, 2006a) are geared towards the marking process. All these tools incorporate some form of scoring rubric. This makes the scoring rubric available in the immediate marking context. Frequently-used comments are attached directly to the rubrics and feedback can be related to the categories within the rubric. This makes the task of the marker more efficient and provides the student with more information. With GradeMark and Re:Mark specialised marking symbols can be defined and attached to the essays. Most of the tools automatically create a summary sheet for each student. This means less time consuming manual steps for the marker and more informative content for the

student. A tool like MarkTool allows the marker to suppress the display of comments belonging to specific categories of a scoring rubric. This helps in focusing the attention during the marking on selected aspects. A further element of advanced support is the ability to extract all marking comments across the essays of a whole class, allowing the instructor to quickly assess the suitability of the wording of the comments and to get an insight into the knowledge level of the class. This is especially important if the instructor has not marked the essays by themselves but has called on the help of markers. MarkTool provides this functionality of extracting all marking comments. Re:Mark provides statistical information on the use of the predefined comments, supporting the instructor in identifying weaknesses across the class.

Once the task of assessing the individual assignments is performed the strengths of computer programs in performing repetitive steps should kick in. A computer tool should be available to automatically extract all comments and marks from the essays for a whole class and return these, together with generic information like the scoring sheet, to each individual student. A few new e-learning tools provide these services. GradeBook (Turnitin.com, 2006b) by Turnitin (Turnitin.com, 2006a) is a companion application to GradeMark. It extracts marks that have been assigned in GradeMark and manages these marks for the whole class. Re:Mark has a similar connection to the Blackboard Gradebook tool, making assignment results available on a class basis. WebCTConnect complements WebCT by addressing the weaknesses of this LMS in the assignment management process. WebCTConnect provides as well a marking environment with scoring rubric and frequently used comments. It can be used by itself or in conjunction with MarkTool. Marked essays, comments and marks can be uploaded to WebCT for a whole class in one operation, saving the instructor the time consuming and error prone steps of sending each marked essay individually to a student. Further, WebCTConnect contains a management feature for group assignments. If the essays have been named according to a predefined rule the tool works out group membership and mark allocation automatically.

E-Learning Support for Marking Teams

The instructor and their marking team need to work together closely as a team to achieve reliable and high quality marking. Even a simple electronic environment

facilitates this teamwork and brings huge advantages compared to a paper approach. General purpose tools like email or shared network drives allow for communication and document exchange. If marked in electronic format it is always easy to retain an unmodified copy of the original essay. Remarketing or parallel marking can therefore be easily achieved. LMS allow setting up teaching assistants or markers giving them access to all submitted assignments.

While generic electronic environments or LMS already bring advantages compared to paper environments there is still plenty of scope for improvement. One issue is that the organisation of the collaboration is left to the human participants. Steps like attaching essays to emails are not difficult but take the team members out of their marking environment and cost time. LMS facilitate shared access to essays but the organisation of which marker assesses which essay is not supported. Two markers can score the same essay but there is no support for the comparison of the marking.

WebCTConnect offers a range of features towards supporting marking teams. It allows the instructor to allocate each marker specific essays for marking. If two markers mark the same essay, e.g., for purposes of inter-marker comparison, both marking results can be stored and compared within WebCTConnect. The tool makes it possible to exchange marking results and marking documents between instructor and markers without leaving the marking environment.

Plagiarism Detection

Plagiarism Prevention (Turnitin.com, 2006d) and SafeAssignment (MyDropBox.com, 2006b) address the issue of plagiarism that has become an increasing problem over the past years. The Internet gives students access to a vast amount of resources that can be used legitimately or illegitimately for essay preparation. The range reaches from properly quoted and referenced material to taking essays from Internet cheat sites. The marker faces the challenge to detect sections in an essay that have been copied and to prove this to the student. Plagiarism detection software provides invaluable support to the marker. It searches resources available on the Internet and on special assignment repositories for matches to an essay. The software highlights these matches in the essay, provides links to the resources copied and calculates a percentage figure for suspected plagiarism. What is left to do for the marker is to decide on the legitimacy of the overlap, that is proper quoting or simple copying, and

then deal with the assignment accordingly. Plagiarism detection tools are used both for educating students and for catching cheaters.

Summary of Advantages of E-Learning Tools for Essay Marking

Although current computing technology cannot perform the core task of actual essay marking, the potential for using computers to assist instructors and markers is enormous. These people can be given guidance and can perform their marking tasks more efficiently. Computing power can make tasks feasible that otherwise are too time consuming to perform. Theoretically, most of the computer tasks described here can be performed manually, either with just general purpose software or even with paper-based systems. Many of the individual steps seem small. Yet, without technology to support extensive commenting, frequent exchange between instructor and markers, and comprehensive analysis of strengths and weaknesses across the whole class these tasks are simply too time consuming. Marking effort is too often reduced to the bare minimum or essay assessment is even replaced with less suitable assessment formats. The use of efficient essay marking support tools can lead to the application of all the marking processes suggested in the educational literature, increasing the reliability and ultimately the quality of essay marking. The more specialised the e-learning support for essay marking the less time consuming work-arounds have to be performed and the more guidance can be integrated, following sound educational principles.

4.4 Tools Provided as Part of Research Projects

This section describes in more detail some specialist tools that provide support for lecturers marking essays. These are essentially home-grown projects developed by enthusiastic staff at learning institutions.

Description of WebCTConnect

WebCTConnect (Massey University, 2006b) offers a variety of functions around assignment management, marking and marking communication. As the name suggests WebCTConnect works in conjunction with the learning management system WebCT. The teacher uses the assignment tool in WebCT to specify assignments. The students

submit their assignments to WebCT and WebCT handles bookkeeping and storage. WebCTConnect provides advanced management facilities. In terms of document management this means the handling of assignments that consist of multiple files and an efficient return of marked assignments to students. All assignment submissions of one class are displayed to the marker in a table view. This table can be sorted by various criteria, for example name of the student or time submitted. A comment, visible only to the marker, can be entered for each assignment. Markers can be allocated to specific assignments and multiple markers can assess the same assignment independently. The makers can exchange their marking comments electronically supported by the tool and the teacher responsible can decide which comments to return to the student. The teacher can develop a scoring rubric, called a marking scheme. For each individual assignment the markers can provide feedback under the various categories of the marking scheme. Under each category the markers can collect a list of frequently used comments that can be copied and adjusted for individual assignments. Comments for each assignment and the marking categories are collated in a summary sheet that is returned to the student. All marking comments across all assignment submissions can be exported into a spreadsheet file. This file can be flexibly configured to contain numeric marks and/or textual feedback at detailed or aggregate levels.

In summary, WebCTConnect supports lecturers with the following features: downloading class lists, assignment definitions and assignments to the lecturer's computer; managing group assignments and marker allocations; setting up of a marking scheme; commenting on assignments in relationship to the marking scheme; exchanging of marking information between markers and lecturer; creating of a summary sheet for each student; efficient upload of marks, comments and commented assignments to WebCT.

Description of MarkTool

MarkTool (Massey University, 2006a) focuses on the actual marking of essays. MarkTool itself does not address assignment management issues. For essays that are submitted as assignments for a whole class of students the management facilities of WebCT and WebCTConnect can be used. If just a small number of individual essays is judged a formal management process might not be required and essays can be

exchanged via general methods like email attachments. The strength of MarkTool lies in the way it allows the marker to attach individual comments directly to the pages of an essay. To start with, the teacher defines a scoring rubric, again called marking scheme. In extension to the marking scheme of WebCTConnect the categories of the MarkTool marking scheme are colour coded with colours defined by the teacher. While marking the teacher then places individual comments on specific pages of an essay. The teacher can identify the exact reference point for a comment by drawing a graphical component, like a rectangle or an ellipse, onto the electronic essay page. Via its colour code the comment is linked to a category of the marking scheme. The teacher has access to various display formats of the marked essay. Marking comments belonging to specific marking categories can be hidden from view. There is a summary page that lists all comments sorted by marking category or marker. Similar to the functionality of WebCTConnect frequently used comments can be developed and marking results across a grouping of essays can be exported into a spreadsheet file. The original essay file is not modified so multiple markers can assess the same essay and can exchange their marking comments. The student receives an electronic copy of their essay that is annotated with individual, colour coded marking comments. Further, the student has access to the marking scheme and a hyperlinked summary page that provides them with an overview of all their comments in relationship to the marking scheme.

In summary, MarkTool supports lecturers with the following: definition of a marking scheme; commenting on assignments with graphical reference areas; relating of comments to marking criteria; maintaining of lists of frequently-used-comments; summary sheet with all comments; separation of comments and marks; extraction of all comments across all assignments for a task.

Analysis of WebCTConnect

As outlined in the section on essay assessment theories earlier in this article a scoring rubric lies at the heart of successful essay marking. WebCTConnect allows the teacher to define such a rubric. WebCTConnect focuses primarily on analytic rubrics yet holistic rubrics can be defined as well. To define an analytic rubric, different marking categories are named, samples solutions or key solution items are described and a numeric mark, the upper limit the student can achieve in this category, is

assigned. For a holistic rubric, the various quality levels are specified and described and the related marks are stated. The rubric needs to be defined before the marking of the essays starts. The rubric can be exported into a separate file and be given to students, or can be sent to markers, for discussion and shared use.

WebCTConnect offers various ways to support the process of marking. As electronic copies of essays are marked and the actual marking data are stored separately from the essay files remarking by the same marker or parallel marking by multiple markers are always possible. WebCTConnect displays all assignments for a class in a table format for easy overview. The markers can add comments to each assignment that will not form part of the feedback given to the student but assist the markers in organising the marking process. That means, a marker can write a note, indicating the status of the marking. This can be used to manage the sequence of the marking, e.g., each essay question at a time, or to convey information like 'review again' or 'check with co-marker'. If holistic marking is used the common sorting of essays into 'piles' can be simulated via the comments. As all assignments are presented in table format the display sequence of the assignments can be changed by sorting via the various column headings. This presents a convenient way of modifying the marking sequence for multiple passes through the assignments.

Assessment theory states that ideally an assignment would be marked by multiple markers. The tertiary reality shows that the lecturer responsible for the assessment is often supported by a team of markers to cope with the workload in large classes. This means that a team of markers needs to collaborate. WebCTConnect facilitates collaboration of marking teams. The leader of the team allocates markers to assignments and distributes the scoring rubric. All members of the team can exchange their marking data. These features form the technical foundations for working towards consistency and quality assurance.

As the theories state individualised feedback is very important for facilitating the understanding and learning of students. With WebCTConnect the markers can write comments for each individual assignment. These comments are aligned to the sections of the scoring rubrics making it easy to give targeted feedback. The student is presented with a summary sheet that shows the details of the scoring rubric, all individual comments and the marks achieved. The writing of individual comments is very time consuming. To assist the marker, frequently used comments can be

collected for each section of a marking rubric. The marker can select from these comments and individualise them for each particular assignment.

An important side-effect of providing feedback to students should be the knowledge the teacher gains on the strengths and weaknesses of a class as a whole. WebCTConnect supports the exporting of all marking comments and marks across all assignments of a whole class into a spreadsheet file. Looking at this file allows the teacher to analyse the marking data according to a variety of criteria. For example, the teacher can look at all comments made under a specific category of a scoring rubric. This will form the basis for gaining information on the level of knowledge of the class as a whole in a particular topic area and can flow into remedial teaching, if required. The distribution of comments and marks across the categories of the scoring rubric can provide insights into the quality of that rubric.

Analysis of MarkTool

As MarkTool can be used in conjunction with WebCTConnect many of the process-based advantages, like assigning of markers, keeping status information and modifying the marking sequence, apply as well. Used without the management support of WebCTConnect some advantages still apply and some new opportunities arise. For example, as electronic copies of essays are marked and the marking data are stored separately, the same essay can still be marked by multiple markers and the results can be compared. Without the management support the submission details of an assignment, especially the name of the student, are not immediately visible. Assuming the students have been instructed to not write their names into their essays, it is possible to fulfil the requirement of anonymous marking.

Like WebCTConnect, MarkTool allows the teacher to setup a scoring rubric. Again, both analytic and holistic scoring rubrics can be defined. Current work on an extension to MarkTool will provide for holistic scoring rubrics in matrix format. A standard holistic scoring rubric defines quality levels for the whole essay task. A more extensive holistic scoring rubric focuses on specific learning targets within the essay task and lists quality levels for each of these targets. Such a holistic scoring rubric will provide better guidance for the markers and will allow assessing different aspects of a piece of work separately. Like in WebCTConnect, all marking data for a group of

assignments can be exported and analysed by the teacher to gain insights for re-teaching.

MarkTool does have several specific strengths. With MarkTool very detailed feedback can be given for each essay. Educational theory advocates for comments being placed as close as possible beside their reference points in the student's work (Renkl & Atkinson, 2002). In MarkTool feedback is placed directly to its reference point in the essay. This provides the direct context between the content of the essay and the comment. The marker can make detailed comments that complement the higher-level comments that can be given to summarise achievement. Because the marker can create graphical reference points for each comment no effort is required to explicitly describe the context of the comments. Additionally, the marker is supported by a frequently used comments mechanism. For the student this means that they can look at each page of their essay and will find localised comments, providing them with detailed, contextualised feedback. As these comments are linked to the categories of the scoring rubric defined in MarkTool, the student gains additional information, guiding them to their strengths and weaknesses in particular areas of their work.

Assessment theory identifies the danger that the marker of an essay can get distracted from assessing the desired learning outcomes of the essay by, for example, elements of presentation. To address this, the use of scoring rubrics is suggested. Scoring rubrics clarify which learning outcomes are targeted and guide the marker throughout the marking process. MarkTool works with scoring rubrics and adds specific support for the marker. As outlined in the previous paragraph, the marker places individual comments that are linked to the categories of the scoring rubric, directly on the pages of the essay. MarkTool offers the marker to choose, for which of the categories comments are visible. If, for example, only one category is chosen, only the comments related to this category will be displayed on the assignment pages. This will allow the marker to focus their attention on the specific category and therefore learning outcome. In a similar fashion, the summary page listing all comments for an essay can be sorted by categories.

Summary of the Contributions of WebCTConnect and MarkTool

Ultimately, the two applications WebCTConnect and MarkTool target improvements for learning and teaching around the essay assessment. For the foreseeable future human markers will be required for the setting of essay tasks and for the assessment of students' efforts. The role of e-learning technology and tools has to be supporting humans to fulfil their tasks more efficiently and at a higher level of quality. The educational theories on essay assessment suggest a range of techniques that so far are hardly supported by e-learning applications. WebCTConnect and MarkTool address a number of these techniques.

Scoring rubrics are at the centre of reliable and efficient essay marking. In both applications the marking is based on such rubrics. WebCTConnect provides extensive support for the management of the marking process and the coordination of marking teams. MarkTool combines detailed individual feedback, which is set in direct context of an essay, with categories inside the scoring rubrics. Both applications provide access to all marking information on class-level and therefore facilitate feedback into teaching. Steps towards quality assurance in team marking are made possible via the exchange of marking data for essays.

Description of Denton's Electronic Feedback Software

Denton (2001; , 2003) developed a system based on Microsoft Word and Excel that allows lecturers to create and email feedback to students. The lecturer sets up the software with the student and assignment details. The feedback is entered into a bank of comments that the lecturer can apply to groups of students. These banks of comments can be setup before the marking begins or as the marking proceeds. The lecturer marks the assignment and a marking sheet is generated with the student name and details, the overall mark and comments. This marking sheet is then emailed to the student.

The feedback is grouped into four categories:

- Grade comments are given according to the students' grade. Once the grades are entered a comment is automatically entered into the marking sheet. This could be excellent work for those who receive marks higher than 80%.

- General comments are sent to the whole class and give overall comments about the assignments.
- Standard comments are given to groups of students such as those who show common misunderstandings. The comments could be based on criteria that are used to mark the essay.
- Personal comments are for individual students.

Analysis of Denton's Electronic Feedback Software

There are some sound benefits to lecturers who use this software. The software design allows lecturers to structure their feedback using good educational practice by encouraging the use of assessment criteria. The system is good at making common comments available for the lecturer and so will save marking time. Most lecturers will know how to use Microsoft Word and Excel. However as Excel has been customised the lecturers will have to learn how to use the system. Finally, the students who receive the feedback suggest that the system should be used more regularly in other courses. Lecturers on other courses have adopted this system.

Denton's system could be made more efficient. The lecturer needs to manually enter in the student details. This step would be avoided if the system could link to the student management system. The system does not provide any way to reference comments to a part of the assignment. Ideally, a comment should be close to the part of the assignment to which it refers. Students can then put the comment into context.

Overall this system is a valuable tool. Student evaluations were positive as were comments from staff who used it. However, widespread uptake of the system did not occur, even with those staff who attended training sessions. This may indicate the difficulties of staff making changes to working practice rather than shortcomings in the software system. While the use of generic software has advantages there are issues with the maintenance required. When the generic software upgrades, changes are also likely in the Denton system.

Description of Markers Assistant

Markers Assistant was designed to improve the management of the assessment process and improve feedback to students (Wells, 2006).

The assignment needs to be setup in Markers Assistant. This includes entering the assignment information, entering the marking guide and adding student information to the system. It is often possible to export student records to a file and to upload them into Markers Assistant. The marking guide allows lecturers to structure their comments within the assessment criteria.

The markers view the assignment in the format it was submitted, e.g. Word, PDF, Excel etc. The students can also submit paper based assignments.

Marks are given using the criteria setup in the structured marking guide. A checkbox helps the marker with this process. The markers comment on misunderstandings with the focus on suggesting solutions to students rather than criticism of the student's work. Examples of good work are also highlighted. The markers comments will go to a moderator before being released to the students. The students receive feedback on a separate feedback sheet that does not link directly to specific sections of the student assignment.

Analysis of the Markers Assistant

Markers Assistant allows for a number of efficiencies over hand-writing on paper. The feedback is stored in a databank so once a comment is entered, the marker can reuse it if it is relevant to other students. Support for marking teams allows markers to work together on a set of assignments. The lecturer co-ordinating the process can moderate the work as the system provides statistics to compare markers and compare the feedback from the markers. Statistics on the marks are provided so the lecturer can identify where the class as a whole needs to do more work.

This system does not allow annotations on the student assignment. Students value comments close to the location to which they refer in the assignment (Renkl & Atkinson, 2002) as this helps them to consider the comment alongside their own attempt.

The markers can sort the essay on student identification or submission date. Some markers may want to sort the essays on other categories such as essay length to identify those who went over or under the word count, and readability scores such as the Flesch Reading Ease score.

Analysis of the Markers Assistant

This is another tool that can help staff provide good quality feedback. There is good documentation and the software is freely available. Markers Assistant encourages good educational practice and will help markers save time.

Description of Classmate

Classmate assists markers by reducing the trivial and repetitive assignment marking tasks (Baillie de Byl, 2004). It is available to staff at the University of Southern Queensland and originated in the Mathematics and Computing Department to assist the marking of a programming course. Classmate focuses on the need to promote effective feedback to large classes where marking teams are used.

Classmate manages student submission, distributes assignments to markers, collates marker feedback and returns the marked assignment to the student. The marker can comment directly into the student assignment with the comments appearing in red. The marker-edited version is a copy of the file the student submitted. The system has good file management which automatically makes copies for markers comments and saves all files for future reference. The marker creates a feedback sheet with marks based on the marking criteria, an overall grade and overall comments. The system then returns to the student the annotated assignment with the feedback sheet.

Analysis of Classmate

The students and staff who have used Classmate are generally positive about it (Baillie-de Byl, 2004). The system provides good educational practice by integrating the marking schedule into the heart of the system and saves staff time. It provides some features that would be difficult to do without a computerised system, such as the management of marking teams.

There are some ways to improve this system. By linking it to other computer systems such as student records, learning management systems and e-learning systems more tasks could be automated. The system is only available to staff in the University of Southern Queensland.

Summary of Tools Provided as Part of Research Projects

The tools reviewed here that have arisen from research projects incorporate educational theories very well. They facilitate high quality assessment and allow teachers and markers to perform their work efficiently. They automate aspects of the assessment work that are of repetitive nature and do not require human intellect. They assist the human marker in the areas where human input is essential and cannot be performed by a computer tool. The biggest problem with these tools is that they are not widely used, not even at the institutions where they have been developed.

5 Issues Concerning the Stakeholders Involved in Assessment

The main stakeholders involved in education are the students, teaching staff and the teaching institutions. Each of these stakeholder groups has its own complex set of expectations, needs and constraints. The following sections highlight important issues concerning the stakeholders in relationship to assessment in general and the use of e-learning for assessment in particular.

5.1 Students

Students raise two major concerns regarding assessment. The first of these concerns relates to the fairness of assessment (Campton & Young, 2005; Nesbit & Burton, 2006). Students question the fairness of assessment. A negative perception of fairness in turn has a negative impact on performance. Students who receive lower marks are more likely to have justice concerns. These concerns reduce motivation, which in turn can have a negative impact on future performance. An improvement on students' views on the fairness of assessment will increase performance (Nesbit & Burton, 2006). The use of e-learning tools might help to make assessment more transparent to students (Jones, Cranston, Behrens, & Jamieson, 2005). This can cause assessment to be perceived as more fair, breaking the negative performance cycle.

The second major student concern relates to feedback (Campton & Young, 2005; Orsmond, Merry, & Reiling, 2005). Students are unsatisfied with the level of feedback they receive (Campton & Young, 2005). They regard a number of feedback comments as being too vague. They criticise receiving too many single comments, supportive or critical, without being given an overall picture (Orsmond, Merry, & Reiling, 2005). A further issue is that students do not know how to take advantage of feedback. They ask for opportunities for dialog with tutors to discuss learning from feedback (Orsmond, Merry, & Reiling, 2005). This is an area where the strengths of e-learning tools in facilitating communication could be used to advantage fairly easily.

Students appreciate opportunities for self-assessment. Multiple-choice tests can be used formatively for self-assessment. The advantage of such tests is that they can be evaluated automatically and provide students with instant response on correctness of

answers. An approach suggested by Peat and Franklin (2002) presents students with short-answer questions and separate answer sheets for self-checking. According to the authors students enjoy the use of this approach. It gives them a different representation of the study material which they use as a learning tool and for self-assessment before exams. O'Reilly (2005) reports that students use self-assessment material to maximise understanding and, given the opportunity, to maximise marks by repeating tests or quizzes. From a staff workload perspective it is only possible to allow multiple submission if e-learning tools provide support. Full automation, encompassing handling and checking of answers, is possible for multiple-choice style tests. While this is not the case for essay-type assessment well-thought out e-learning support for submission, management and return of student work would provide an important step towards facilitating more opportunities for learning around assessment material.

Struyven et al. (2005) discuss student preferences between course work and examinations. They name multiple choice tests and essays as the two main assessment forms for examinations. According to their research the majority of students prefer multiple choice tests as they perceive these as easier, allowing them to choose among options that contain the correct answer already. This provides a chance of success by luck and does not require students to formulate answers by themselves. Yet, students with good learning skills and high confidence in their academic abilities prefer essay examinations. Given the choice between examinations and coursework Struyven et al. (2005) report that students prefer coursework and here specifically newer forms of assessment methods like portfolios and peer- and self-assessment. Students appreciate that these forms of assessment encourage understanding instead of focusing on memorising. They regard these forms of assessment as fairer as they provide ongoing assessment instead of taking a snapshot on one day (Struyven, Dochy, & Janssens, 2005). E-learning approaches provide a strong opportunity for supporting such ongoing forms of assessment, in particular in the areas of document management and tracking and facilitation of communication among teachers, students and peers.

Naturally, students are very sensitive about assessment. If e-systems are used for assessment students expect these systems to function without failure (Blayney & Freeman, 2004; Freeman & McKenzie, 2002). They are unforgiving towards any problems, especially if these relate to loss of data. One strength of web-based systems

is their around-the-clock availability. Students appreciate and have come to expect this availability and do not tolerate any downtime (Freeman & McKenzie, 2002).

The use of a web-based system for assessment should not be an add-on. It should be part of an environment students use anyway for other tasks related to their studies, otherwise acceptance of this assessment system is low (Freeman & McKenzie, 2002). This is particularly important for low-stakes or purely formative assessments, as students, juggling various commitments, tend to put their effort where they see the biggest short-term gain. A well-integrated assessment system will lower the cost of entry and increase the chances of participation.

Orsmond et al. (2005) report that a number of students read and reread comments. The students keep their marked assignments so that they are able to refer back to the feedback provided to them by their teachers. Electronic documents can be conveniently stored and electronic copies of marked assignments can facilitate students in referring back to previous work. While this is possible in principle with standard file storage systems a lot more could be done is assisting students to create annotated repositories of marked work.

A very practical issue relates to the readability of handwritten comments. Students have difficulties in deciphering the handwritten comments put on their work (Blayney & Freeman, 2004; Bridge & Appleyard, 2005; Higgins, Hartley, & Skelton, 2002). Typed comments are easier to read and, if looked at on screen, have the additional advantages of various display sizes and of searching and sorting. What is required to make it practical for the marker to type comments are suitable marking systems. Such systems need to facilitate commenting on multiple levels, 'inline', on specific issues and as summary. As reported earlier students ask for feedback that addresses their work in totality. Additionally, students request comments on specific issues, placed directly on their work in the page margins (Orsmond, Merry, & Reiling, 2005).

5.2 Teaching Staff

A number of different groups of staff are involved in the assessment process. There are the academics or teachers who are responsible for the design of the assessment, including choice of the assessment format, links to learning outcomes and overall responsibility. Tutors play a role in preparing students for assessment, conducting

assessment steps and responding to post-assessment questions from students. Markers, often employed on a casual basis, work with marking guidelines to assess student work summatively and provide feedback. Administrative staff deal with the handling of student work, in receiving and returning copies, and keeping track of submission details and marks. Technical staff can be involved in the setup and monitoring of assessment systems. In many cases one individual will fulfil several or even all of the roles described. The following paragraphs use the generic term 'staff' to refer to the people involved in the assessment process.

The use of e-systems for assessment can increase the awareness of staff regarding assessment approaches (Aller et al., 2005) and can positively influence the link between assessment and teaching and learning (Buzzetto-More & Alade, 2006). Aller and colleagues describe the use of a web-based assessment library that guides staff through the selection of adaptable assessment instruments appropriate to their needs. The system provides staff with clear assessment expectations and models of communication and teamwork activities. An important side-effect of using the system has been increased discussion among staff about assessment methods. The authors further report that the web-based system has resulted in staff providing clearer and more specific feedback.

E-systems for assessment can lead to time savings and efficiency gains. Campton and Young (2005) report on a semi-automated assessment system that uses scoring rubrics and some automated comments. The authors claim that at least 25% of time for commenting was saved using this system. Blayney and Freeman (2004) have developed an approach of using spreadsheets with built-in formulae for self-assessment in business subjects. Tasks with individualised parameters can be generated for students who can self-assess their solutions before submitting their work for teacher assessment. This system provides learning opportunities for students and efficiency gains to staff. Jones and colleagues (2005) report on the development and use of an online assignment submission and management systems. Some staff have reported a considerable time saving for marking when using the system. According to the researchers a crucial factor in the assessment of the effectiveness of the system lies in the flexibility of staff. Staff who are willing to adjust their assessment practice can benefit greatly from such an electronic system. These approaches reported on here

provide three quite different examples on using technology to increase efficiency in assessing student work that clearly goes beyond the multiple-choice approach.

Earlier sections of this review have emphasised that an assessment approach must be chosen for pedagogical reasons with the learning outcomes in mind. In similar fashion the selection of e-learning tools for assessment must be guided by the pedagogical design of the assessment. It is up to staff to do the thinking, take control and match tool use to their assessment approach and learning goals (Freeman & McKenzie, 2002). The primary focus needs to be on pedagogical improvements while productivity gains can only be hoped for in the medium to long term (Warburton, 2006). Further, it is responsibility of staff to ensure that students recognise the link between tool use and learning goals and receive appropriate guidance. O'Reilly (2005) reports on exemplar approaches of online assessment. The author examined the factors contributing to the success of these approaches and identified staff commitment to pedagogical rationales for adoption as most significant. She goes on to describe the characteristics of what she calls 'hallmarks of excellence' in online assessment as the blending of on- and off-campus cohorts, formative feedback before summative marks, reduction and streamlining of the marking workload and authentic assessment tasks.

Essential to the successful adoption of e-systems for assessment and e-learning tools in general is a strong support network for staff. There needs to be good support among all staff involved in the assessment, including technical support personnel, as well as general support from colleagues and departments (Freeman & McKenzie, 2002). It has been reported that staff often experience disinterest and a lack of support from colleagues (O'Reilly, 2005). Ill-informed opinions and rumours on problems with system performance and capability heard from colleagues form barriers to the adoption of e-systems for assessment (Jones, Cranston, Behrens, & Jamieson, 2005). Staff interests as compared to institutional needs have to be at the centre of the decision to adopt an e-learning system (Jones, Cranston, Behrens, & Jamieson, 2005; O'Reilly, 2005). Staff capabilities, technical knowledge and experiences with computer tools play a further important role in the decisions about uptake (Blayney & Freeman, 2004; Jones, Cranston, Behrens, & Jamieson, 2005). It is essential that suitable e-systems are available that are efficient and easy to use. A complete approach must be offered that not only covers assignment submission and

management but as well the actual marking process (Jones, Cranston, Behrens, & Jamieson, 2005). If this is not the case staff are faced with the decision of either to print student work and mark on paper or to use unsuitable tools for the marking of the electronic assignment copies.

The use of e-systems to support essay-type and other formative assessment approaches is not yet wide-spread. Approaches are being developed and trialled by early adopters (Sim, Holifield, & Brown, 2004). These are staff who are actively seeking to improve teaching and learning via assessment. O'Reilly (2005) has analysed her exemplar approaches of online assessment against the model of scholarship of teaching proposed by Trigwell et al. (Trigwell, Martin, Benjamin, & Prosser, 2000). This model captures both the scholarship of teaching and teaching as scholarship. The model expresses levels of engagement as staff being informed about teaching and learning issues, applying this knowledge in their own teaching practice, targeting improvements in student learning and disseminating their findings on teaching and learning to the benefit of the wider community of students and colleagues. O'Reilly has found clear links to the model and therefore to a research approach to teaching and higher education in all the exemplars studied. This is a very positive finding as it shows that new approaches with a solid scholarly foundation targeting improvements in teaching and learning are being developed.

The main idea behind formative assessment is to improve student learning. To achieve this staff must give useful and appropriate feedback to students. Learning how to write such feedback is a difficult task. Most institutions of higher learning offer courses on teaching and learning issues to staff. While these courses cover assessment in general a wide variety of assessment-related topics must be discussed, leaving only little room for formative feedback. Additionally, such a complex topic cannot be fully covered in a one-off training session. Ongoing attention is required. Staff are learning while 'doing'. In this context McKenzie (McKenzie, 2004) proposes the interesting idea of using e-systems to support this learning process. In an e-system it is possible to provide a marking team with access to each other's marking comments. This allows newer staff to learn from more experienced colleagues and has the additional benefit of this learning occurring in a discipline specific context.

5.3 Institutions

Institutional support is crucial for the successful adoption of e-learning tools for assessment (Buzzetto-More & Alade, 2006; Freeman & McKenzie, 2002; Sim, Holifield, & Brown, 2004; Warburton, 2006). The introduction of a new assessment programme requires planning and foresight. Institutional and administrative support, preparation for the technological challenges and long-term planning for sustainability are required (Buzzetto-More & Alade, 2006). Institutional support needs to include the environment surrounding staff. Particularly, support is required from colleagues, departments and technical staff (Freeman & McKenzie, 2002). Uptake should be motivated not mandated and promoted both horizontally among colleagues as well as top-down by learning technologists. A coordinated strategy is required that includes buy-in by senior management and consideration of external influences like government initiatives (Warburton, 2006). An institutional strategy is important to guide departments in the selection of appropriate approaches and tools. Without such strategy individual departments will choose varying systems, causing a range of negative consequences. Students will have to cope with different interfaces, there will be increased licence costs, and higher demands on administrative and technical support (Sim, Holifield, & Brown, 2004). Institutions need to create a support network of experienced colleagues and learning technologists. Institutions need to put supportive procedures in place, show long-term commitment, provide secure funding and a strong infrastructure (Warburton, 2006).

Kenny (2002) reports on the introduction of a university-wide e-learning environment at his university. Many challenges had to be overcome in the selection of appropriate hardware and software, the development of efficient and secure systems around these choices and the integration of these systems with existing administrative and student access infrastructure. By setting institutional targets for uptake and providing initial training for staff a rapid uptake of the e-learning environment by staff was achieved. Following on from this the important point Kenny makes is the necessity of looking beyond simple uptake to the creation of meaningful learning experiences. Achieving this requires changes in teaching practices and the redevelopment and reorganisation of learning materials. This in turn is only possible with the professional support of educational design and production staff. Multi-skilled development teams need to work closely with teaching staff. On an institutional level this has direct implications

for staff workloads and resource planning. Warburton (2006) raises similar points in context of the introduction of computer assisted assessment. Important is the focus on pedagogical improvements. Productivity gains can only be hoped for in the medium-to long-term. Warburton emphasises the care that has to be taken in the introduction of assessment-related systems as compared to general e-learning systems, based on the credit-bearing nature of the former. O'Reilly (2005) writes that the fundamental principles of good assessment must always remain at the forefront of considerations when introducing any form of e-supported assessment.

The creation of meaningful learning experiences and the consideration of fundamental principles of good assessment are related closely to the quality of assessment. Students have expressed criticisms on the quality of their learning experiences in context of the use of e-learning systems (Weaver, Nair, & Spratt, 2005). A lack of quality was related back to situations where staff had not volunteered to use such systems but had followed institutional requirements. A number of conclusions can be drawn from these observations. Institutions need to encourage staff to take ownership of their use of e-learning systems and they need to provide better support. Further, institutions need to monitor quality and not just quantity of uptake. Buzzetto-More and Alade (2006) advocate for the establishment of review committees on institutional level. Tasks of such committees would include the development, implementation and review of assessment plans. The impact on student learning and performance had to be monitored and results fed back into future actions. While these authors emphasise the impact on students Warburton (2006) points to the other stakeholders involved when evaluating the success of an e-assessment approach. Tutors will be motivated by time savings and will consider it as important to cause no unrest among students. Learning technologists will focus on the use of advanced learning technologies for sound pedagogical purposes, while institutions might look for spread of uptake and the integration with other systems.

The introduction of e-assessment systems requires a holistic institutional approach (Downton, 2006; Kennedy, Webster, Benson, James, & Bailey, 2002; Kenny, 2002). All systems used at an institution need to be integrated, with common access points for serving staff, student and institutional needs (Kennedy, Webster, Benson, James, & Bailey, 2002). Assessment facilities are highly important both for staff and students and are included among the top five main functionalities for both groups (Kennedy,

Webster, Benson, James, & Bailey, 2002). An example for a holistic approach can be shown with the integration of an assignment submission and management system with a student management system (Downton, 2006). If integrated student class lists will be seamlessly transferred into the assignment system and the enrolment status of each student will be maintained correctly. Such coordination is well within the capability of computer systems and interfaces between systems have to be implemented accordingly. These issues are important as neglect in this area would increase staff workload unnecessarily.

Selecting the most appropriate e-learning environments is a difficult undertaking (Kenny, 2002; Lindblom-Ylanne & Pihlajamaki, 2003). One issue is the specific knowledge that is required, on technical, instructional and institutional levels. Teaching staff in general will not possess this knowledge. The other issues relates to the complex network of stakeholders and systems in place in higher education institutions. Teaching staff cannot work in isolation and their system choices will affect others, students, colleagues and administrators. As outlined earlier every e-learning or assessment system needs to function in context of other computer systems. For these reasons institutions must take a leading role in the selection and implementation of e-learning environments.

Regardless which set of systems is selected problems will occur. Especially in the sensitive area of assessment sound procedures need to be in place to deal with such problems in efficient and effective ways (Pain & Le Heron, 2003). Preparation is required to limit any potential damage and facilitate recovery from difficulties. It is important to know responsibilities and pathways to resolving the problems. In the assessment area the crucial issue is credibility, both for students and staff.

6 Importance of E-Learning Technology for Assessment

E-learning technology already has a very important role in teaching and learning. Assessment of student learning and evaluation of instruction are of critical importance and can and need to be supported by e-learning technology and strategies (Buzzetto-More & Alade, 2006). Good assessment, especially of formative nature, is complex and substantial effort is required in this area. E-learning technology makes high-quality formative assessment practical by removing some of the constraints limiting higher uptake (Committee on the Foundations of Assessment, 2001). E-learning technologies and approaches can assist in the assessment cycle for formative assessment of student work (Blayney & Freeman, 2004). In brief, this cycle contains stages for the exchange of documents, for communication, for producing artefacts addressing the assessment tasks and for writing feedback. E-learning technologies can assist in all these stages.

There are a variety of ways in which e-learning can facilitate assessment in general and essay-type assessment in particular. Freeman and McKenzie (2002) advocate for the use of web-based systems to improve students' learning of team working skills. They state that such systems reduce problems with teamwork, especially in the context of large classes. Stevens and Jamieson (2002) talk about the use of e-tools to improve marking quality and feedback. Blayney and Freeman (2004) list a number of ways in which computing support can help with providing feedback. They talk about supporting human markers in commenting on supply items, about polling student opinions in the classroom and about getting insights into student understanding with quizzes and test. Plimmer and Mason (Plimmer & Mason, 2006) as well as Edwards and colleagues (2002) emphasise the advantages of electronic submission and handling of assignments. They list issues like the ease in collection of student work, the removal of geographic limitations, the reduced risk of lost work, the time and resource savings if printing is not required, the long-term availability based on the ease of storage of electronic artefacts, and the efficient return of marked student work. A number of authors point out the opportunities for anonymous participation and marking that arise from e-assessment approaches (Davies, 2002; Downton, 2006; Edwards, Fernandez, Milionis, & Williamson, 2002). Student identities can be stored in a database and the artefact that the marker sees can be kept anonymous, which is

one for the recommendations for addressing the challenges in the reliability of essay marking (Gronlund, 2006; Lambert & Lines, 2000). Peer environments can be designed in which students can participate anonymously and only the assessing teacher knows the students identity. New assessment designs are possible in which students learn from repositories of marked example assignments (Heinrich, 2004; Heinrich & Lu, 2005). These designs will include self- and peer-assessment and scaffolding of student learning by only partially releasing feedback data.

It is essential to use technology for the right pedagogical reasons. The use of technology for its own sake will not improve educational assessment (Committee on the Foundations of Assessment, 2001). The use of technology must be well thought-out and assessment must be integrated into learning (Blayney & Freeman, 2004). Students should be informed of the reasons for using e-learning systems and the relation of these systems to learning design (Freeman & McKenzie, 2002). New approaches to assessment in higher education can be informed by e-learning technology. These approaches should emphasise the role of technology in supporting the human assessor, whose role is vital, especially for formative assessment (McGuire, 2005). For any form of assessment validity and reliability are important measures. In an e-learning context new forms of assessment can and should be explored to achieve these measures and this process for improvement should be driven by all stakeholders involved in education (Wijekumar, Ferguson, & Wagoner, 2006).

O'Reilly (2005) names various aspects of learning designs that become possible and manageable by utilising e-learning technologies. New forms of assessment can be integrated into these learning designs. Online archives can be constructed and used for reflection, critique and accountability. On-campus and off-campus students can join discussions via asynchronous dialogue tools with the exchange of perspectives benefiting both cohorts. The formative and summative assessment of assignments can be decoupled. Instead of just submitting a final version students incrementally improve their work. Along the way they receive formative feedback, from staff or peers. Such a design requires consistent commitment from students and minimises the risks of plagiarism and cheating. Self-assessment can play an important role in increasing understanding and recognising weaknesses. O'Reilly acknowledges the role of automated feedback as can be derived from quizzes. Time saving for staff can

be achieved by using these automated methods of feedback on restricted response items but as well by using tools that support onscreen marking.

In summary it can be said that there is lots of opportunity for using e-learning technologies and approaches for improvements in assessment. There is also consensus in the literature that these opportunities are not been taking advantage of and that not enough effort is put into exploring the use of e-learning for assessment processes and tasks (Baillie-de Byl, 2004; Blayney & Freeman, 2004; O'Reilly, 2005; Wells, 2006).

7 Summary and Preliminary Conclusions

In this review the term essay-type assessment has been defined as the formative assessment of coursework where students supply items such as essays and other forms of documents. The importance of this type of assessment lies in the challenges it poses to students to formulate their own thoughts and construct their own answers, linking to higher level learning outcomes. The educational value of such assessment has been established and the parameters determining high quality assessment have been determined. The review has highlighted areas in which e-learning tools and approaches can support essay-type assessment. The main areas are the management of assignment submission, storage and return, the assistance in providing individualised feedback to students, the coordination of marking teams and the communication with students.

Various types of software can be used to support essay-type assessment. Learning management systems provide base level support, largely for the management of assignments. General purpose tools like word processing systems can be adapted for tasks like providing feedback. Specialised assessment software is available. It falls into three major groupings of which only one is relevant to essay-type assessment. Specialist software is available for the specification and automated marking of restricted response item assessment like multiple choice tests. These tools are not relevant in the context of this research as they target a different assessment paradigm. Some specialist software attempts the automated assessment of free-form student writing. Several articles in the literature report a relatively good reliability of these systems. The main reasons why these systems are not considered further in this research are the need for very large essay sample sizes to calibrate the automated

marking, restrictions in providing meaningful individualised feedback to students and the limitations to text-only student work. The last group of specialist software for assessment is comprised of tools for the marking of assignments. These tools build on educational assessment theories. They offer support for the management of assignments and, most importantly, provide features conducive to high quality formative marking. The tools integrate scoring rubrics with feedback on detailed and summary level. The general philosophy of these tools is to liberate the marker from the mundane but necessary administrative tasks and to create an environment that facilitates quality feedback enhancing student learning. While the software in this last group shows a lot of promise for essay-type assessment it is not widely used. Uptake is localised to institutions where researchers have developed these tools and even at these institutions only early adopters are involved.

This review presents the main groups of stakeholders concerned with assessment as students, staff and institutions. Students in general are concerned with the fairness of assessment and the level of feedback they receive. Students are open to new forms of assessment and to the use of e-learning tools in general. Based on the high stakes involved in assessment students are not prepared to tolerate any form of malfunction of an assessment system. Staff live in a complex network of their own, student and institutional needs. Their considerations regarding uptake of e-learning tools for assessment are influenced by opinions and experiences of colleagues, familiarity with technology, opportunities for efficiency gains, preparedness in modifying their approaches to teaching and learning and institutional support. The last point of institutional support is of critical importance. Institutions must be instrumental in making the appropriate e-learning systems available, guiding and protecting staff with suitable policies, providing comprehensive training both on instructional and technical levels and ensuring ongoing support.

The preliminary conclusions that are drawn from this literature review are as follows.

- Essay-type assessment is of highest educational value.
- Strong opportunities exist for using e-learning technologies, tools and approaches to facilitate essay-type assessment.
- E-learning technologies, used for the appropriate pedagogical reasons, show a high potential for improvements in the assessment area.

- General purpose software and current learning management systems provide limited support for essay-type assessment.
- Specialised software, available in localised initiatives, has been developed that incorporates educational theories and efficiently facilitates high-quality essay-type assessment.
- The uptake of both general and specialised software for supporting essay-type assessment is limited.
- Institutions need to provide comprehensive and ongoing instructional and technical support to move from early adopters to general uptake of e-learning techniques, tools and approaches for essay-type assessment.

The section on policy documents written by the tertiary providers has shown substantial differences between universities and Institutes of Technologies and Polytechnics (ITPs). The university documents address both procedural issues and the role of assessment for learning. The documents of the ITPs focus solely on procedural issues. Only some of the literature outside the policy documents reported on assessment examples at specific institutions. Overall, no information could be gained to distinguish assessment practices by institution type. It will be interesting to see if the further investigations in the course of this project will shed more light on this area.

References

- Adobe. (2006). Adobe Acrobat Professional. Retrieved 05/10/2006, from <http://www.adobe.com/products/acrobatpro/>
- Ala-Mutka, K., Uimonen, T., & Järvinen, H.-M. (2004). Courses with Automatic Program Style Assessment. *Journal of Information Technology Education, 3*, 245-262.
- Alam, L. S. (2004). Is Plagiarism More Prevalent in Some Forms of Assessment Than Others? *21st ASCILITE Conference*, 48-57.
- Aller, B. M., Kline, A. A., Tsang, E., Aravamuthan, R., Rasmusson, A. C., & Phillips, C. (2005). WeBAL: a web-based assessment library to enhance teaching and learning in engineering. *IEEE Transactions on Education, 48*(4), 764-771.
- Assessment Reform Group. (1999). *Assessment for learning: beyond the black box*. Cambridge, U.K: University of Cambridge School of Education.
- Assessment Reform Group. (2002). *Testing, motivation and learning*. Cambridge, U.K: University of Cambridge School of Education.
- Auckland University of Technology. (2002). Guidelines On The Policy For Assessment Of Student Achievement [Electronic Version]. Retrieved 25/10/2005 from http://www.aut.ac.nz/staff/academic_quality_office/policies_guidelines_notes.htm.
- Baillie-de Byl, P. (2004). An Online Assistant for Remote, Distributed Critiquing of Electronically Submitted Assessment. *Educational Technology & Society, 7*(1), 29-41.
- Barrett, C., & Luca, J. (2002). *Open online assessment: keeping the tutors honest!* Paper presented at the Winds of change in the sea of learning: Charting the course of digital education. Proceedings of the 20th ASCILITE Conference, Auckland, New Zealand.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*, 7-74.
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139-149.
- Blackboard. (2006). Blackboard. from <http://www.blackboard.com/>
- Blayney, P., & Freeman, M. (2004). Automated formative feedback and summative assessment using individualised spreadsheet assignments. *Australasian Journal of Educational Technology, 20*(2), 209-231.
- Bridge, P., & Appleyard, R. (2005). System failure: A comparison of electronic and paper-based assignment submission, marking, and feedback. *British Journal of Educational Technology, 36*(4), 669-671.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 79*, 474-482.
- Buzzetto-More, N. A., & Alade, A. J. (2006). Best Practices in e-Assessment. *Journal of Information Technology Education, 5*, 251-269.
- Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgement is involved: the features of an e-marking tool. *Assessment & Evaluation in Higher Education, 30*(5), 529-537.

- Campton, P., & Young, J. (2005). *Please sir, may I have some more? A comparative study on student satisfaction with assessment feedback methods in an undergraduate unit*. Paper presented at the Balance, Fidelity, Mobility: maintaining the momentum? The 22nd annual conference of the Australasian Society for computers in learning in tertiary education (ascilite), Brisbane, Australia.
- Chen, C.-M., & Chen, Y.-Y. (2005). *Learning Performance Assessment Approach Using Learning Portfolio for E-learning systems*. Paper presented at the ICALT: 5th IEEE International Conference on Advanced Learning Technologies, Kaohsiung, Taiwan.
- Committee on the Foundations of Assessment. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.
- Conlon, T. (2006). Formative Assessment of Classroom Concept Maps: the Reasonable Fallible Analyser. *Journal of Interactive Learning Research*, 17(1), 15-36.
- Creative Technology. (2006). Markin. Retrieved 05/10/2006, from <http://www.cict.co.uk/software/markin/index.htm>
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438-481.
- Crooks, T. J. (2004). Tensions between assessment for learning and assessment for qualifications, *Paper presented at the third conference of the Association of Commonwealth Examinations and Accreditation Bodies*. Nadi, Fiji.
- Davies, P. (2002). Using Student Reflective Self-Assessment for Awarding Degree Classifications. *Innovations in Education and Teaching International*, 39(4), 307-319.
- Davies, P. (2003). *The Automatic Generation of 'Marks for Marking' within the Computerised Peer-Assessment of Essays* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough, U.K.
- de Lambert, K., Ellen, N., & Taylor, L. (2006). Chalkface challenges: a study of academic dishonesty amongst students in New Zealand tertiary institutions. *Assessment & Evaluation in Higher Education*, 31(5), 485-503.
- Denton, P. (2001). *Generating and e-Mailing Feedback to Students Using MS Office*. Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Denton, P. (2003). *Evaluation of the 'Electronic Feedback' Marking Assistant and Analysis of a Novel Collusion Detection Facility* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Dikli, S. (2006). An Overview of Automated Scoring of Essays. *Journal of Technology, Learning, and Assessment*, 5(1).
- Downton, A. (2006). *Online Coursework Submission from Pilot to University-wide Implementation: Rationale, Challenges and Further Development*. Paper presented at the 10th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Edwards, K. I., Fernandez, E., Milionis, T. M., & Williamson, D. M. (2002). EAST: developing an electronic assessment and storage tool. *Assessment & Evaluation in Higher Education*, 27(1), 95-104.

- Freeman, M., & McKenzie, J. (2002). SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *British Journal of Educational Technology*, 33(5), 551-569.
- Ge, X., & Er, N. (2005). An online support system to scaffold real-world problem solving. *Interactive Learning Environments*, 13(3), 139-157.
- Goodfellow, R., & Lea, M. (2005). Supporting writing for assessment in online learning. *Assessment & Evaluation in Higher Education*, 30(3), 261-271.
- Gronlund, N. E. (2006). *Assessment of Student Achievement*. Boston: Pearson.
- Hand, B. M., Prain, V., & Yore, L. (2001). Sequential Writing Tasks' Influence on Science Learning. In L. M. Paivi Tynjala, Kirsti Lonka (Ed.), *Writing as a Learning Tool: Integrating Theory and Practice* (pp. 105 - 129). Dordrecht, Boston, London: Kluwer Academic Publishers.
- Hanna, G. S., & Dettmer, P. A. (2004). *Assessment for Effective Teaching Using Context-Adaptive Planning*. New York: Pearson.
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education*, 10, 169-207.
- Harlen, W., & James, M. (1996). Creating a positive impact of assessment on learning, *Paper presented at the annual meeting of the American Educational Research Association*. New York.
- Hattie, J. (1999). Influences on student learning, *Inaugural professorial lecture*. University of Auckland, New Zealand.
- Heinrich, E. (2004). Electronic Repositories of Marked Student Work. *Journal of Educational Technology & Society*, 7(3), 82-96.
- Heinrich, E., & Lu, J. Y. (2005). Opportunities in learning from marked student work. *ED-Media 2005 World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 4470-4475.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The Conscientious Consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64.
- ITP New Zealand. (2006). Academic Quality Standards [Electronic Version]. Retrieved 10/10/2006 from <http://www.itpq.ac.nz/documents/AcademicAuditStandards2006Revision.pdf>.
- Jafari, A., McGee, P., & Carmean, C. (2006). Managing Courses, Defining Learning: What Faculty, Students, and Administrators Want. *Educause review*(July/August), 50-70.
- Jones, D., Cranston, M., Behrens, S., & Jamieson, K. (2005, 9-11 November). *What makes ICT implementation successful: A case study of online assignment submission*. Paper presented at the Open Learning and Distance Learning Association of Australasia (ODLAA), University of South Australia.
- Kendle, A., & Northcote, M. (2000). *The Struggle for Balance in the Use of Quantitative and Qualitative Online Assessment Tasks*. Paper presented at the The 17th conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE). Coffs Harbour, NSW.
- Kennedy, D. M., Webster, L., Benson, R., James, D., & Bailey, N. (2002). My.monash: Supporting students and staff in teaching, learning and administration. *Australian Journal of Educational Technology*, 18(1), 24-39.
- Kenny, J. (2002). Managing innovation in educational institutions. *Australian Journal of Educational Technology*, 18(3), 359-376.

- Korhonen, A., Malmi, L., Nikander, J., & Tenhunen, P. (2003). Interaction and Feedback in Automatically Assessed Algorithm Simulation Exercises. *Journal of Information Technology Education*, 2, 241-256.
- Kutay, C., & Ho, P. (2005). Designing agents for feedback using the documents produced in learning. *International Journal on E-Learning*, 4(1), 21(18).
- Lambert, D., & Lines, D. (2000). *Understanding Assessment: Purposes, Perceptions, Practice*. Padstow: TJ International.
- Leathwood, C. (2005). Assessment policy and practice in higher education: purpose, standards and equity. *Assessment & Evaluation in Higher Education*, 30(3), 307-324.
- Lindblom-Ylänne, S., & Pihlajamäki, H. (2003). Can a collaborative network environment enhance essay-writing processes? *British Journal of Educational Technology*, 34(1), 17-30.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and Assessment in Teaching*. Columbus: Pearson Merrill Prentice Hall.
- Liu, C.-C., & Tsai, C.-M. (2005). Peer assessment through web-based knowledge acquisition: tools to support conceptual awareness. *Innovations in Education and Teaching International*, 42(1), 43-59.
- Liu, K.-Y., Wan, J.-Y., & Chen, H.-Y. (2005). *A marking-based Synchronised Multimedia Tutoring System for Composition Studies*. Paper presented at the ICALT: 5th IEEE International Conference on Advanced Learning Technologies., Kaohsiung, Taiwan.
- Macdonald, J. (2003). Assessing online collaborative learning: process and product. *Computers & Education*, 40(4), 377-391.
- Macdonald, R., & Carroll, J. (2006). Plagiarism - a complex issue requiring a holistic institutional approach. *Assessment & Evaluation in Higher Education*, 31(2), 233-245.
- Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*, 29(3), 311-321.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. F. Tanner (Ed.), *Critical issues in curriculum, Eighty-seventh yearbook of the National Society for the Study of Education, Part 1* (pp. 83-121). Chicago, IL: University of Chicago Press.
- Mason, O., & Grove-Stephensen, I. (2002). *Automated free text marking with Paperless School*. Paper presented at the 6th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Massey University. (2004). Guidelines for the Preparation and Administration of Examinations and Other Assessments [Electronic Version]. Retrieved 25/10/2005 from <http://policyguide.massey.ac.nz/Procedures/AssessmentExamGuide.doc>.
- Massey University. (2006a). eLearning Support for Formative Assessment - MarkTool. Retrieved 05/10/2006, from <http://www-ist.massey.ac.nz/marktool/>
- Massey University. (2006b). eLearning Support for Formative Assessment - WebCTConnect. Retrieved 05/10/2006, from <http://www-ist.massey.ac.nz/marktool/>
- McGuire, L. (2005). Assessment using new technology. *Innovations in Education and Teaching International*, 42(3), 265-276.
- McKeever, L. (2006). Online plagiarism detection services: saviour or scourge? *Assessment & Evaluation in Higher Education*, 31(2), 155-165.

- McKenzie, S. (2004, 5-8 December). *Assessing quality of feedback in online marking databases: An opportunity for academic professional development or just Big Brother?* Paper presented at the Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference, Perth.
- McLachlan-Smith, C., & Irons, B. (1998). *Ideas to Share: Examples of Successful Extramural Study Guide Design*. Palmerston North: Massey University.
- Moodle. (2006). Moodle. from <http://moodle.org>
- Moreale, E., Whitelock, D., Raw, Y., & Watt, S. (2002). *What Measures do we Need to Build an Electronic Monitoring Tool for Postgraduate Tutor Marked Assignments?* Paper presented at the 6th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Muirhead, B. (Artist). (2002). *Effective Online Assessment Strategies for Today's Colleges & Universities*
- MyDropBox.com. (2006a). Re:Mark. Retrieved 05/10/2006, from <http://www.mydropbox.com/services/safeassignment.php>
- MyDropBox.com. (2006b). SafeAssignment Plagiarism Prevention. Retrieved 05/10/2006, from <http://www.mydropbox.com/services/safeassignment.php>
- Nelson Marlborough Institute of Technology. (2005). Moderation of Assessment [Electronic Version]. Retrieved 10/10/2006 from <http://www.nmit.ac.nz/FutureStudent/Library/2I2%20Moderation%20of%20Assessment.pdf>.
- Nesbit, P., & Burton, S. (2006). Student justice perceptions following assignment feedback. *Assessment & Evaluation in Higher Education*, 31(6), 655-670.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Nitko, A. J. (2004a). *Educational Assessment of Students* (4th ed.). Upper Saddle River, New Jersey: Pearson Education.
- Nitko, A. J. (2004b). *Educational Assessment Outcomes*. Columbus: Pearson Merrill Prentice Hall.
- O'Reilly, M. (2005, 13th - 15th June). *Hallmarks of Excellence in Online Assessment*, . Paper presented at the The First International Conference on Enhancing Teaching and Learning Through Assessment, Hong Kong Polytechnic University.
- Orsmond, P., Merry, S., & Reiling, K. (2005). Biology student's utilization of tutor's formative feedback: a qualitative interview study. *Assessment & Evaluation in Higher Education*, 30(4), 369-386.
- Pain, D., & Le Heron, J. (2003). WebCT and Online Assessment: The best thing since SOAP? *Educational Technology & Society*, 6(2), 62-71.
- Parameswaran, A., & Devi, P. (2006). Student plagiarism and faculty responsibility in undergraduate engineering labs. *Higher Education Research and Development*, 25(3), 263-276.
- Peat, M., & Franklin, S. (2002). Supporting student learning: the use of computer-based formative assessment modules. *British Journal of Educational Technology*, 33(5), 515-523.
- Plimmer, B., & Mason, P. (2006). *A Pen-based Paperless Environment for Annotating and Marking Student Assignments*. Paper presented at the Proc. Seventh Australasian User Interface Conference (AUIC2006), Hobart, Australia.

- Renkl, A., & Atkinson, R. K. (2002). Learning From Examples: Fostering Self-Explanations in Computer-Based Learning Environments. *Interactive Learning Environments*, 10(2), 105-119.
- Rudner, L. M., Garcia, V., & Welch, C. (2006). An Evaluation of the IntelliMetricSM Essay Scoring System. *The Journal of Technology, Learning, and Assessment*, 4(4).
- Sadler, D. R. (1987). The specification and promulgation of achievement standards. *Oxford Review of Education*.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Sim, G., Holifield, P., & Brown, M. (2004). Implementation of Computer-assisted Assessment: Lessons from the Literature. *ALT-J, Research in Learning Technology*., 12(3), 215-229.
- Sim, G., Read, J. C., & Holifield, P. (2006). *Evaluating the User Experience in CAA Environments: What Affects User satisfaction?*
- . Paper presented at the 10th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Stevens, K., & Jamieson, R. (2002). The Introduction and Assessment of Three Teaching Tools (WebCT, Mindtrail, EVE) into a Post Graduate Course. *Journal of Information Technology Education*, 1(4), 233-252.
- Stiggins, R. J. (2005). *Student-Involved Assessment for Learning* (4th Edition ed.). Upper Saddle River: Pearson Merrill Prentice Hall.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Student's perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341.
- Sung, Y.-T., Chang, K.-E., Chiou, S.-K., & Hou, H.-T. (2005). The design and application of a web-based self- and peer-assessment system. *Computers & Education*, 45(2), 187-202.
- Tai Poutini Polytechnic. (2006). Quality Management System - Policies [Electronic Version]. Retrieved 10/10/2006 from <http://www.tpp.ac.nz/pdfs/tpp-qms-2006.pdf>.
- Taras, M. (2002). Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education*, 27(6), 501-510.
- The Open Polytechnic. (2006). Assessment and Moderation Regulations [Electronic Version]. Retrieved 10/10/2006 from [http://www.openpolytechnic.ac.nz/studentinformation/regulationsandpolicies/2006 assessment and moderation regs.pdf](http://www.openpolytechnic.ac.nz/studentinformation/regulationsandpolicies/2006%20assessment%20and%20moderation%20regs.pdf).
- Torrance, H., & Pryor, J. (1998). *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Buckingham: Open University Press.
- Trigwell, K., Martin, E., Benjamin, J., & Prosser, M. (2000). Scholarship of Teaching: a model. *Higher Education Research and Development*, 19(2), 155-168.
- Turnitin.com. (2006a). Effective Learning Tools. Retrieved 05/10/2006, from <http://www.turnitin.com/static/index.html>
- Turnitin.com. (2006b). GradeBook. Retrieved 05/10/2006, from <http://www.turnitin.com/static/index.html>
- Turnitin.com. (2006c). GradeMark. Retrieved 05/10/2006, from <http://www.turnitin.com/static/index.html>

- Turnitin.com. (2006d). Plagiarism Prevention. Retrieved 05/10/2006, from <http://www.turnitin.com/static/index.html>
- Tynjala, P., Mason, L., & Lonka, K. (2001). Writing as a learning tool: an introduction. In L. M. Paivi Tynjala, Kirsti Lonka (Ed.), *Writing as a Learning Tool: Integrating Theory and Practice* (Vol. 7, pp. 7-22). Dordrecht, Boston, London: Kluwer Academic Publishers.
- Universal College of Learning. (2006). Assessment and Moderation Policy.
- University of Auckland. (2002). Teaching and Learning Policy: Assessment of Student Learning [Electronic Version]. Retrieved 25/10/2005 from <http://www.auckland.ac.nz/uoafms/default/uoafms/about/teaching/policiesprocedures/docs/studentlearning.pdf>.
- University of Canterbury. (2004). Assessment Guidelines [Electronic Version]. Retrieved 25/10/2005 from <http://www.canterbury.ac.nz/ucpolicy/index.aspx>.
- University of Otago. (2002). The Teaching and Learning Plan [Electronic Version]. Retrieved 25/10/2005 from <http://www.otago.ac.nz/about/pdfs/teachinglearningplan.pdf>.
- Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An Overview of Current Research on Automated Essay Grading. *Journal of Information Technology Education*, 2.
- Vastani, H., Edwards, S. H., & Pérez-Quñones, M. A. (2005). *Supporting On-line Direct Markup and Evaluation of Students' Projects*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition.
- Waiariki Institute of Technology. (2006). Academic Statute [Electronic Version]. Retrieved 10/10/2006 from <http://pmf.waiariki.ac.nz/Wit%20Planning%20&%20Quality%20Management/Statute%20&%20Delegations/Academic%20Statute.doc>.
- Waikato Institute of Technology. (2006). Academic Regulations [Electronic Version]. Retrieved 10/10/2006 from <http://www.wintec.ac.nz/files/mywintec/policies/acregscurrent.pdf>.
- Warburton, B. (2006). *Quick Win or Slow Burn? Modelling UK HE CAA Uptake* Paper presented at the 10th International Computer Assisted Assessment Conference, Loughborough, U.K.
- Weaver, D., Nair, C. S., & Spratt, C. (2005). *Evaluation: WebCT and the Student Experience*. Paper presented at the Making a Difference: 2005 Evaluations and Assessment Conference, Sydney.
- WebCT. (2006). WebCT. from <http://webct.com/>
- Wells, J. (2006). Markers Assistant--a software solution for the management of the assessment process. *International Journal on E-Learning*, 5(3), 439(420).
- Wijekumar, K., Ferguson, L., & Wagoner, D. (2006). Problems with Assessment Validity and Reliability in Web-Based Distance Learning Environments and Solutions. *Journal of Educational Multimedia and Hypermedia*, 15(2), 199-215.
- Wong, C. K., Wong, W., & Yeung, C. H. (2001). Student Behaviour and Performance in Using a Web-Based Assessment System. *Innovations in Education and Teaching International*, 38(4), 339-346.
- Wu, Y.-F. B., & Chen, X. (2005). *eLearning Assessment Through Textual Analysis of Class Discussions*. Paper presented at the ICALT: 5th IEEE International Conference on Advanced Learning Technologies., Kaohsiung, Taiwan.

- Zhang, J., & Heinrich, E. (2005a). A system designed to support formative assessment of open-ended written assignments. *5th International Conference on Advanced Learning Technologies, ICALT*, 88-92.
- Zhang, J., & Heinrich, E. (2005b). Using computers to support formative assessment of assignments. *ED-Media 2005 World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 4510-4515.

Appendix A – Sources for Literature Review

Journals

AJEL - Australasian Journal of Educational Technology
Alt-J - Association for Learning Technology Journal
Assessment & Evaluation in Higher Education
BJET - British Journal of Educational Technology
CIHB - Computers in Human Behavior
CJLT - Canadian Journal of Learning and Technology
Computers & Education
Education and Information Technologies
EDUCAUSE Quarterly (EQ)
ETR&D - Educational Technology Research & Development
ETS - Journal of Educational Technology & Society
HERD - Higher Education Research and Development
IEEE Transactions on Education
IETI - Innovations in Education and Teaching International
IJEEE - International Journal of Electrical Engineering Education
IJEL - International Journal on E-Learning
ILE - Interactive Learning Environments
IMEJ - Interactive Multimedia Electronic Journal of Computer-Enhanced Learning
Innovative Higher Education
Interacting with Computers
International Journal of Educational Research
International Journal of Technology and Design Education
IRRODL - The International Review of Research in Open and Distance Learning
JCAL - Journal of Computer Assisted Learning
JEE - Journal of Engineering Education
JEMH - Journal of Educational Multimedia and Hypermedia
JILR - Journal of Interactive Learning Research
Journal of Distance Education
Journal of Distance Learning DEANZ
Journal of Information Technology Education jite.org
Journal of Technology, Learning and Assessment
JRTE - Journal of Research on Technology in Education (Former Journal of Research on Computing in Education)
LL&T - Language Learning & Technology
TICL - Technology, Instruction, Cognition and Learning

Conference Proceedings

CAA conference proceedings
ASCILITE conferences proceedings
HERDSA Annual Conference Proceedings

Online Databases

Ed/IT Lib Education & Information Technology Library (formerly AACE Digital Library)

International Conference on Mathematics / Science Education and Technology
Society for Information Technology and Teacher Education International
Conference
World Conference on E-Learning in Corporate, Government, Healthcare, and
Higher Education
World Conference on Educational Multimedia, Hypermedia and
Telecommunications
AACE Journal
Contemporary Issues in Technology and Teacher Education
Information Technology in Childhood Education Annual
International Journal of Educational Telecommunications
International Journal on E-Learning
Journal of Computers in Mathematics and Science Teaching
Journal of Educational Multimedia and Hypermedia
Journal of Interactive Learning Research
Journal of Technology and Teacher Education
WebNet Journal: Internet Technologies, Applications & Issues

EDUCAUSE

EDUCAUSE Quarterly (EQ)

EDUCAUSE Review

Additional resources contributed by: EDUCAUSE Committees, EDUCAUSE
Learning Initiative, EDUCAUSE Policy Office, Security Task Force,
ECAR, EDUCAUSE Office, Forum for the Future of Higher Education,
Web Seminars by EDUCAUSE, CAUSE/EFFECT Archives, Educom
Review Archives

Books on Education

Bowen-Clewley, L., & Strachan, J. (1997). *A Report on an Assessment Stocktake Undertaken for the New Zealand Qualifications Authority*.

Brown, S., Race, P., & Bull, J. (1999). *Computer-Assisted Assessment in Higher Education*. London: Clays.

Chatterji, M. (2003). *Designing and Using Tools for Educational Assessment*. Boston: Allyn & Bacon.

Committee on the Foundations of Assessment. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

Cooper(Ed), P. (1983). *Towards Internal Assessment*. Wellington: Technical Correspondence Institute.

Council, N. R. (2001). *Knowing what Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

Ecclestone, K. (1994). *Understanding Assessment*. Leicester: NIACE.

Edwards, A., & (Eds), P. K. (1995). *Assessing Competence in Higher Education*. London: Biddles.

- Flanagan, D. P., & Harrison(Eds), P. L. (2005). *Contemporary Intellectual Assessment*. New York: Guilford.
- Freeman, R., & Lewis, R. (1998). *Planning and Implementing Assessment*. London: Biddle.
- Gipps, C. V. (1994). *Beyond Testing: Towards a Theory of Educational Assessment*. London: The Falmer Press.
- Griffin, P., & Nix, P. (1991). *Educational Assessment and Reporting: A New approach*. Melbourne: Harcourt Brace Jovanovich.
- Gronlund, N. E. (1988). *How to Make Achievement tests and Assessments*. (Fifth edition ed.). Boston: Allyn& Bacon.
- Gronlund, N. E. (2006). *Assessment of Student Achievement*. Boston: Pearson.
- Hand, B. M., Prain, V., & Yore, L. (2001). Sequential Writing Tasks' Infulence on Science Learning. In L. M. Paivi Tynjala, Kirsti Lonka (Ed.), *Writing as a Learning Tool: Integrating Theory and Practice* (pp. 105 - 129). Dordrecht, Boston, London: Kluwer Academic Publishers.
- Hanna, G. S., & Dettmer, P. A. (2004). *Assessment for Effective Teaching Using Context-Adaptive Planning*. New York: Pearson.
- Hart, D. (1994). *Authentic Assessment: A Handbook for Educators*. Parsippany, New Jersey: Dale Seymour.
- Kane, M. B., & (Eds), R. M. (1996). *Implementing Performance Assessment: Promises, Problems and Challenges*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Kubiszyn, T., & Borich, G. (2000). *Educational Testing and Measurement: Classroom Application and Practice*. New York: Wiley & Sons.
- Lambert, D., & Lines, D. (2000). *Understanding Assessment: Purposes, Perceptions, Practice*. Padstow: TJ International.
- Linn, R. L. (1993). *Educational Measurement* (Third Edition ed.). Phoenix: Oryx Press.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and Assessment in Teaching*. Columbus: Pearson Merrill Prentice Hall.
- Marzano, R. J., Pickering, D., & McTighe, J. (1994). *Assessing Student Outcomes*. Alexandria: Association for Supervision and Curriculum Development.
- Ministry of Education. (1999). *The Use of Exemplars in Outcomes-based Curricula: An International Review of the Literature*. Auckland: Auckland Uniservices Ltd.
- National Society for the Study of Education. (2004). *Towards Coherence Between Classroom Assessment and Accountability*. Chicago: University of Chicago Press.
- New Zealand Qualifications Authority. (1997). *To Your Marks! Advice to teachers and tutors on setting and marking assessments*. New Zealand: NZQA.
- Nitko, A. J. (2004). *Educational Assessment of Students* (4th ed.). Upper Saddle River, New Jersey: Pearson Education.
- Nitko, A. J. (2004). *Educational Assessment Outcomes*. Columbus: Pearson Merrill Prentice Hall.
- Paivi Tynjala, L. M., Kirsti Lonka (Ed.). (2001b). *Writing as a Learning Tool: Integrating Theory and Practice* (Vol. 7). Dordrecht, Boston, London: Kluwer Academic Publishers.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical Considerations in Computer-Based Testing*. New York: Springer-Verlag.

- Quinn, S. L., & Kanter, S. B. (1982). *How to Pass an Essay Examination*. Dubuque: Kendall.
- Reynolds, C. R., Livingston, R. B., & Wilson, V. (2006). *Measurement and Assessment in Education*. Boston: Pearson Education.
- Stiggins, R. J. (2005). *Student-Involved Assessment for Learning* (4th Edition ed.). Upper Saddle River: Pearson Merrill Prentice Hall.
- Tanner, D. E. (2001). *Assessing Academic Achievement*. Boston: Allyn & Bacon.
- Tanner, H., & Jones, S. (2003). *Marking and Assessment*. New York: Continuum.
- Torrance, H., & Pryor, J. (1998). *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Buckingham: Open University Press.
- Walker, C., & Schmidt, E. (2004). *Smart Tests: Teacher-made tests that help students learn*. Markham: Pembroke.
- Westerheijden, D. F., Brennan, J., & (Eds), P. A. M. M. (1994). *Changing Contexts of Quality Assessment*. Utrecht: Lemma.
- Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.
- Williams, D. D., Howell, S. L., & Hricko, M. (2006). *Online Assessment, Measurement and Evaluation*. London: Information Science Publishing.
- Wilson, L. W. (2005). *What Every Teacher Needs to Know about Assessment*. Poughkeepsie: Eye on Education.