Addressing Efficiency and Quality of Marking in Essay Assessment with E-learning Support

EVA HEINRICH
INSTITUTE OF INFORMATION SCIENCE AND TECHNOLOGY, MASSEY UNIVERSITY, PALMERSTON NORTH, NEW ZEALAND

INTRODUCTION Assessment can be divided into two major categories that are often applied in combination: formative assessment, often called assessment for learning, and summative assessment, often called assessment of learning. Formative assessment focuses on providing learners with feedback to assist them in recognising their knowledge gaps and offering guidance for closing these gaps, while summative assessment concentrates on assessing the current level of knowledge without the focus on guidance of learning (Black & Wiliam, 1998). Formative assessment is regarded as being of very high educational value (Black & Wiliam, 1998; Sadler, 1998; Yorke, 2003).

A prominent format for formative assessment is the essay format. 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, 2004). Essays aim at the highest level of learning outcomes by emphasising the integration and application of ideas (Gronlund, 2006).

The importance of formative assessment and assignments in essay format is reflected in tertiary practise and policy. The policy documents of New Zealand universities (Auckland University of Technology, 2002; Massey University, 2004; Massey University, 2005; University of Auckland, 2002; University of Canterbury, 2004; University of Otago, 2002; University of Waikato, 2005) emphasise the need for a mixture of coursework, combining summative and formative aspects, and summative final examinations. Assessment that encourages independent thinking and deep learning and emphasises the development of communication, research, and thinking skills is favoured. The policies suggest that students need to be assisted in their learning and given guidance towards improvement by receiving a clear indication of the strengths and weaknesses presented in their work.

One of the practical challenges around assessment with essays 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 are not possible. Therefore, the success of e-learning tools in supporting essay assessment has been very limited. Today’s widely used learning management systems, WebCT (2005), Blackboard (2005), and Moodle (2005), provide basic support around the management of essay assignments. The systems allow the setting up of assignment parameters, regulate student access, deal with submissions, and administer summative marks. More advanced tasks, like the allocation of markers to individual essays or the exchange of marking information among markers, are not handled by these tools. Most critically, no support is provided for the actual marking task.

The next section of this article explores assessment theory around essays in more detail. Then two novel e-learning tools for essay assessment are introduced and grounded in the theory.

**ASSESSMENT THEORY & ESSAYS**

**Clarification of Essays**

Essays can be classified as “supply items” (Gronlund, 2006). In contrast to selection-type assessment items such as multiple-choice tests, essays require students to retrieve material from texts or memory and to compose answers in their own words (Hanna & Dettmer, 2004). Essays are characterised by freedom of response that is not available when only selecting from predefined choices. Writing an essay demands the ability to create, organise, and integrate data and ideas (Linn & Miller, 2005). Whereas selection-type assessment encourages students to remember and interpret the ideas of others, essays focus on the highest-level learning outcomes of analysing, evaluating, and creating (Gronlund, 2006).

**Advantages of Writing**

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). Writing, as required in formulating essays, plays an important part in achieving higher-level learning outcomes.

**Essays Across Disciplines**

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. In the natural sciences writing has proven to be beneficial for the understanding of scientific concepts (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.

**Forms of Essay Questions**

The assessment literature distinguishes two forms of essay questions, namely restricted response questions and extended response questions (Gronlund, 1988; Hanna & Dettmer, 2004; Linn & Miller, 2005; Nitko, 2004). 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 length. Extended response questions might pose some practical limitations, such 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.

**Marking of Essays**

The nature of essay 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, 2004). To address these challenges the literature presents a very coherent picture of suggested techniques and procedures (Gronlund, 1988, 2006; Hanna & Dettmer, 2004; Lambert & Lines, 2000; Linn & Miller, 2005; Nitko, 2004).

**Scoring Rubrics**

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 rubric supporting either analytic or 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. With extended response questions and the inherent degree of freedom they provide, it is not possible to develop 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 in the right direction.

**The Marking Process**

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. Focussing on each question separately implies the marker has to work through the essays multiple 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 marking.

**Individualised Feedback**

Returning to the formative aspects of essay assessment, the importance of feedback needs to be emphasised. The marker should provide feedback to each student, outlining strengths and weaknesses in their work thus guiding them towards further learning (Linn & Miller, 2005; Nitko, 2004; 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, 2004). While facilitating student learning is the most important aspect of feedback, a further advantage is the conclusions that can be drawn for teaching. By 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, 2004). 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).

**E-LEARNING TOOLS FOR ESSAY-TYPE ASSESSMENT:**

**WEBCTCONNECT AND MARKTOOL**

The previous section of this article has discussed theoretical and research aspects of formative assessment using essays and has given a brief insight into the status of e-learning support for essay assignments. Simple assignment management tasks are facilitated by current learning management systems, yet no support is provided for the actual marking. Faced with this situation and the importance of essays for higher-level learning, two novel e-learning applications for essay assignments have been developed. The following paragraphs will first briefly introduce these applications from a functional perspective and will then describe how they are based on the educational theories outlined earlier.

**WebCTConnect Functional Description**

The application WebCTConnect offers a variety of functions around assignment management, marking, and marking communication. As suggested by its name, 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 it handles assignment management 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. 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 markers can exchange their marking comments and the teacher responsible can decide which comments to return to the student. The teacher can develop a scoring rubric, called 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.

**MarkTool Functional Description**

The application MarkTool focuses on the actual marking of essays. MarkTool itself does not address 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 involved, a formal management process might not be required and essays can be exchanged via general methods like e-mail attachments. The strength of MarkTool lies in the way it allows the marker to attach individual comments directly to the pages of an essay in electronic form.

To start with, the teacher defines a scoring rubric or marking scheme. 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 WebCT Connect, 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 which is annotated with individual colour-coded marking comments. Further, the student has access to the marking scheme and a hyperlinked summary page which provides them with an overview of all their comments in relationship to the marking scheme.

Table 1 presents an overview showing how WebCT, WebCTConnect, and MarkTool can be combined. A detailed description of the applications, links to related research publications, and software downloads are available at the Web page http://www-ist.massey.ac.nz/MarkTool/.
Table 1 The combination of WebCT, WebCTConnect, and MarkTool for the marking of essay-type assignments

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lecturer defines assignment parameters in WebCT</td>
</tr>
<tr>
<td>2</td>
<td>Students submit assignments to WebCT</td>
</tr>
<tr>
<td>3</td>
<td>Lecturer downloads submission information (date, time) and assignment files using WebCTConnect</td>
</tr>
<tr>
<td>4</td>
<td>Lecturer assigns markers using WebCTConnect</td>
</tr>
<tr>
<td>5</td>
<td>Lecturer/markers assess assignments using MarkTool, providing detailed feedback relating to marking categories</td>
</tr>
<tr>
<td>6</td>
<td>Lecturer and markers communicate about marking using WebCTConnect</td>
</tr>
<tr>
<td>7</td>
<td>Lecturer analyses feedback across all assignments using MarkTool</td>
</tr>
<tr>
<td>8</td>
<td>Lecturer uploads marks and marking feedback to WebCT</td>
</tr>
<tr>
<td>9</td>
<td>Students access WebCT to see their marks and marking feedback</td>
</tr>
</tbody>
</table>

Analysis of WebCTConnect in Light of Assessment Theories

A scoring rubric lies at the heart of successful essay marking. WebCT Connect 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, sample 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 it 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 is always possible. WebCTConnect displays all assignments for a class in a table format for easy overview. The markers can add comments to each assignment which will not form part of the feedback given to the student but will 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.

The ideal in essay marking is that
each 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.

Individualised feedback is very
important for facilitating student
understanding and learning. With
WebCTConnect the markers can write
comments for each individual assign-
ment. 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 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 in
Light of Assessment Theories
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 set up 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. As in WebCT Connect, 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, with comments being placed as close as possible beside their reference points in the student’s work (Renkl & Atkinson, 2002). In MarkTool feedback is linked 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 they can look at each page of their essay and 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 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 allows 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 allows 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 WebCT Connect 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 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.

A BRIEF LOOK AT OTHER E-LEARNING APPROACHES FOR ESSAY ASSESSMENT

Generic computing tools can be used to simulate some aspects of essay marking. For example, word processing or PDF annotation programs can be used to add feedback to essays. The disadvantage of such 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 the 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. What suffers is the quality of marking, and opportunities for learning and teaching are missed.

The assignment tools of all common learning management systems only provide limited assignment management support and no assistance for the actual marking of essays. It can only be hoped that the developers of these systems will add new functionalities in the near future.

On the positive side, there are some tools available that address specific issues in essay marking. Turnitin (2006) offers a set of tools for essay marking. At the core is the detection of plagiarisms. Turnitin compares essays against a huge database of previously submitted essays and information available on the Internet. Overlaps are noted and percentages of “sameness” are calculated. This assists the marker in deciding if a student has plagiarised. Using the GradeMark tool, the marker attaches comments to an essay and assigns marks. GradeMark contains a simple form of an analytic scoring rubric and comments can be displayed as icons directly on the essays. The GradeBook tool facilitates the management of essay marks for a whole class.

The Markin (2006) software allows the marker to insert either predefined symbols or textual comments into the essay text. Additionally, summary comments can be given. The tool calculates some statistical information around the number and type of comments made. Assignment management is addressed by supporting the marker in returning marked essays to students via e-mail.

CONCLUSIONS & FUTURE WORK

The value and importance of formative assessment via essays is emphasised strongly in the educational literature. The nature of essay assessment brings with
it challenges for reliable and valid marking of essays. The literature addresses these challenges by proposing the careful design of essay questions, the use of suitable scoring rubrics, and the development of appropriate management procedures for marking. The currently widely used e-learning environments provide only very limited support for essay assessment. Two novel applications, WebCTConnect and MarkTool, have been designed to fill this gap in the e-learning landscape. These applications incorporate the principles of good essay marking as suggested in the literature. WebCTConnect focuses on the management procedures for assignment marking, whereas MarkTool targets the provision of detailed feedback closely linked to a scoring rubric. Overall, the applications aim at facilitating efficient, high-quality marking that supports learning and teaching. The hope is that the availability of such tools will help to lower the barriers for essay assessment and contribute to higher-quality marking outcomes.

While the availability of specialised tools is an important step in the right direction, it is by itself not sufficient. One issue is the need for integrating such tools with common e-learning platforms. Many of the features derived from the assessment literature and implemented in WebCTConnect and MarkTool could be implemented in today’s widely used learning management systems. The developers of these systems have so far largely focused on summative assessment and have given only little attention to general assignment management issues. Improvements in these systems towards formative assessment with essays, as outlined in this article, would be a big step forward towards high-quality assessment in support of learning.

A second issue lies with the education of teachers, lecturers, and markers. At tertiary level the instructors are often subject experts with limited formal knowledge of learning and assessment theories. Training courses that are offered as part of professional development often stay on a fairly abstract level. This is particularly true for formative assessment. Hands-on courses linked to tool support are available for summative assessment, yet the same does not apply in the formative area. Practical and specific information on how to set up a scoring rubric or on how to formulate feedback that contributes to learning is rarely presented. This situation could change with the availability of e-learning tools for essay marking. Repositories of scoring rubrics and marking comments could be developed and integrated into training systems. These repositories could be annotated by education specialists in regard to the educational soundness of the examples contained. Sorted by subject areas this would provide an excellent resource of concrete examples to guide academics in their professional development.

REFERENCES


---

Eva Heinrich is a Senior Lecturer in Computer Science at Massey University. Her main research interests are in e-learning support for formative assessment and in e-portfolios for computer science and engineering students. Heinrich leads the research project “E-learning Framework for Formative Assessment,” funded under the eCDF scheme of the New Zealand Tertiary Education Commission.