Open Research Online

The Open University's repository of research publications and other research outputs

Advice for Action with Automatic Feedback Systems

Book Section

How to cite:

Whitelock, Denise (2018). Advice for Action with Automatic Feedback Systems. In: Caballé, Santi and Conesa, Jordi eds. Software Data Engineering for Network eLearning Environments. Lecture Notes on Data Engineering and Communications Technologies, 11. Springer, pp. 139–160.

For guidance on citations see [FAQs](#).

© 2018 Springer International Publishing AG

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

http://dx.doi.org/doi:10.1007/978-3-319-68318-8_7

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Advice for Action with Automatic Feedback Systems

Denise Whitelock

The Open University, UK

Abstract

This Chapter reviews the role of feedback in supporting student learning. It highlights some of the problems that persist with providing meaningful feedback, which should preferably take the form of providing advice, that can be actioned by the student. It then discusses the progress made with automatic feedback through a number of case studies which include the OpenEssayist, Open Comment and OpenMentor computer assisted feedback systems.

Findings suggest feedback that provides socio-emotive support to students, together with recognising their effort, in turn encourages the student to continue working on a problem. The use of automatic hints also moves the feedback closer to "Advice for Action". Building tools with automatic feedback to support both students and tutors can relieve some of the continual pressure on staff resources and three case studies are presented below that address this issue.

Keywords: Automatic Feedback, Advice for Action, Socio-cognitive feedback, draft essays feedback

1. Introduction

There is a general recognition in the Assessment field that times are changing and that assessment needs to become embedded in the teaching/learning cycle, and not purely as a checking device for the awarding institution. e-Assessment has tried to address this issue by providing timely and constructive feedback to students through the development of a number of interactive tasks that can be automatically marked, often presented to the student in a multiple choice question format but more importantly can provide immediate feedback to the learner. There is also a recognition that assessment tasks themselves must change and that feedback too merits a reconceptualization (Merry et al, 2013).

2. The role of feedback in teaching and learning

Feedback is a common feature of educational practice (e.g. Black & Wiliam, 1998), and one that has been widely researched but not necessarily implemented or understood to its full potential in practice. This has led to a large amount of research attempting to define what feedback is, when it should be used, and how it could be made more beneficial for students and tutors. Beaumont, O'Doherty & Shannon (2011) for instance identify the "fundamental aim of feedback practice, which is to progressively and explicitly develop students' self-evaluative skills through engagement in the process" (p.683). From this we can see that feedback should have the intention not just of reporting back on finished work, but also of offering advice to self-motivated learners on where they can improve in future work.

As Evans (2013) explained, "Even when "good" feedback has been given, the gap between receiving and acting on feedback can be wide given the complexity of how students make sense of, use, and give feedback (Taras, 2003)" (p.94). Therefore feedback needs to be viewed by tutors and students as an ongoing activity within the cycle of course learning, which feeds into further learning, rather than as an add-on or end point of summative assessment. This is the concept that other researchers have referred to as "feed-forward" (Evans, 2013; Hattie & Timperley, 2007). Thus feedback must be presented in a way that participants can understand, and that they can interpret in terms of where improvements can be made in the future. Hattie & Timperley argued that feedback must be a follow-up to information given to learners, so that they are aware of task requirements before their work is judged by their teacher.

Therefore, feedback is a central part of the teaching and learning process, but one that must follow task instruction and be followed by space for reflection and scope to implement suggestions. Narciss (2013) progressed this notion through classifying the functions of feedback as cognitive, metacognitive and motivational. It has been proposed that 'mediators' (especially 'understanding feedback' and 'agreement with feedback') operate between the provision of feedback features and

implementation of suggestions (Nelson & Schunn, 2009). The researchers suggested that cognitive feedback factors were most likely to influence understanding, and affective factors were more likely to influence agreement. If we want students to make use of feedback, it is important in designing course resources to consider how to ensure that feedback is understood. Nelson & Schunn (2009) also claimed that feedback involved motivation, reinforcement and information. These collective functions of feedback may be particularly important for students who are returning to study after a period of time in employment, who may find it more difficult to understand and access Higher Education modes of study (Scott et al., 2011).

In terms of the purpose of feedback, Chickering & Gamson (1987) outlined seven principles of good practice for undergraduate education, of which the third was “encourages active learning”. Likewise, Nicol & Macfarlane-Dick (2006) stated that students should be urged to be proactive rather than reactive with regard to feedback, using it as a springboard for improvement rather than a stop point. Therefore, feedback or tutor input must do more than just identify misconceptions in students’ work. It must motivate learners to engage with the topic and the task, so that their work comes from and demonstrates understanding rather than just doing enough to get a mark.

For some years now, many courses and universities have made increasing use of technology to support assignment delivery and submission, as well as the medium for offering feedback. Learning has become radically more open and self-regulated, as well as hugely evolved with the innovative uses of new technology. As Steffens (2006) highlighted, “In parallel to the rising interest in self-regulation and self-regulated learning, the rapid development of the Information and Communication Technologies (ICT) has made it possible to develop highly sophisticated Technology-Enhanced Learning Environments (TELEs)” (p.353).

Computer-provided feedback and assessment has some way to go to catch up with these innovations, particularly where courses cater for large numbers of students. The ability to offer automated guidance and feedback at the point of student need to large numbers could help to revolutionise the experience and performance of teaching and learning in higher education. This is particularly pertinent as many universities, including the institution where the study reported in this paper took place, are increasingly catering for distance and round- the-clock learners, many of whom are out of the practice of academic writing.

Feedback involves motivation, reinforcement and information (Nelson & Schunn, 2009). The researchers addressed five features of feedback: summarization; specificity; explanations; scope (local or global); and affective language. We have drawn on these five features in determining the types of feedback to offer on students’ draft essays. Referring to the first feature, summarization, it was claimed that ‘receiving summaries has previously been found to benefit performance: when college students received summaries about their writing, they made more substantial revisions. . . . Therefore, receiving summaries in feedback is expected to promote more feedback implementations’ (Nelson & Schunn, 2009, p. 378).

Feedback is moving forward at a pace with data analysis playing a prominent role. Data can now be collected unobtrusively and during learning activities. However collecting more data does not mean it is necessarily informative and cannot support the teaching and learning dynamic of the classroom. Often it is the feedback from the data analysis that is important and with the development of automatic feedback systems, it is the “Advice for Action” (Whitelock, 2011) that is advocated in this chapter.

3. Feedback problems and progress with Automatic Feedback

There are a number of problems in Higher Education which include progression and retention. Student “dropout” is an even more fundamental problem in distance education. Simpson (2012) has reported that students drop out before they submit their first assignment and this can be as high as a 30% dropout rate for some modules studied at the UK’s Open University. Therefore there is a need to increase students’ confidence and skills during the early weeks of study. The ideal solution would be for students to receive bespoke feedback from their tutors but as this is a resource intensive solution, could automated formative feedback provide a solution?

The following section (4) describes an application known as “Open Essayist” that delivers automated formative feedback designed to help university students improve their assignments. If automated feedback were to be delivered to students to assist them to acquire certain skills, such as essay writing, which will then reduce dropout rates, will the machine feedback always be well received? Is there not an emotional response to feedback that needs to be recognised? Human tutors convey empathy and modulate their feedback to match their students’ immediate response to feedback in face-to-face settings using tone of voice and facial expression. However recognising the students’ contribution and effort is important too. This is a course of action recommended by Mueller & Dweck (1998), where constructive feedback that recognises effort, in turn encourages the student to continue working on a problem and also moves the feedback closer to “Advice for Action”. An automatic feedback system that addresses this particular issue is known as “Open Comment” and is discussed in Section 5.

Another problem with feedback, when it is delivered through an electronic medium, is a lack of recognition of the socio-emotive effect it might evoke in the learner. Too much negativity can be demoralising and again results in student dropout, while a lack of praise can also be demotivating. Teachers need to recognise this issue and receive support in giving feedback that will assist the student emotionally and match the mark that was awarded for an assessment task.

Section 6 describes a system, known as OpenMentor, which analyses and displays the different types of tutor comments provided as feedback to the student. It then provides reflective comments to the tutor about their use of feedback, in order to encourage them to provide a balanced combination of socio-emotive and cognitive support together with ensuring that the feedback is relevant to the assigned grade. Building tools with automatic feedback to support both students and tutors can relieve some of the continual pressure on staff resources and three case studies are presented below that address this issue.

4. OpenEssayist

OpenEssayist (Field et al, 2013) made use of a Natural Language Analytics engine to provide direct feedback to students when they were preparing an essay for summative assessment. The challenge was to provide meaningful feedback to the students themselves so that they could self-correct rather than providing them with a recommender system which elicits a tutor intervention (Arnold & Pistilli, 2012). OpenEssayist provides automated feedback on draft essays, and was developed as part of the SAFeSEA project, to specifically help the UK’s Open University students with their essay writing skills.

OpenEssayist is a real-time learning analytics tool, which operates through the combination of a linguistic analysis engine, which processes the text in the essay, and a Web application that uses the output of the linguistic analysis engine to generate the feedback. OpenEssayist was built because many students at the Open University return to study after some time spent in the workforce, and so it is common that a significant period of time has passed since their last experience of writing academic essays. It is therefore not surprising that many find this task difficult, and without adequate support may decide to leave their course. This is one crucial reason why a system that can intervene and offer support between students’ draft and final submitted essays might be so valuable for students and tutors alike. In creating a system that can go some way to meeting these needs, a number of preliminary studies were made (Alden et al., 2014; Alden, Whitelock, Richardson, Field, & Pulman, 2014). These illustrated that the underlying premises for the construction of OpenEssayist were sound and hence merited further development.

The final system was then developed to process open-text essays and offered feedback through key phrase extraction and extractive summarisation. Key phrase extraction identifies which individual words or short phrases are the most suggestive of an essay’s content, while extractive summarisation essentially identifies whole key sentences. This operates under the assumption that the quality and position of key phrases and key sentences within an essay illustrate how complete and well-structured the essay is, and therefore provide a basis for building suitable models of feedback.

There are a number of ‘automated essay scoring’ (AES) or ‘automated writing evaluation’ (AWE) systems in existence and some are commercially available (see Criterion (Burstein et al., 2003). Pearson’s WriteToLearn (based on Landauer’s Intelligent Essay Assessor (Landauer et al., 2003) and

Summary Street (Franzke & Streeter, 2006), IntelliMetric (Rudner et al., 2006), and LightSIDE (Mayfield & Rose, 2013), all include feedback functionality, though they have their roots in systems designed to attribute a grade to a piece of work. The primary concern of these systems is to help the user make stepwise improvements to a piece of writing. In contrast, the primary concern of OpenEssayist is to promote self-regulated learning, self-knowledge, and metacognition. Rather than telling the user in detail how to fix the incorrect and poor attributes of her essay, OpenEssayist encourages the user to reflect on the content of her essay. It uses linguistic technologies, graphics, animations, and interactive exercises to enable the user to comprehend the content of his/her essay more objectively, and to reflect on whether the essay adequately conveys his/her intended meanings. Writing-Pal (Dai et al., 2011; McNamara et al., 2011) is the system that is most similar to ours in that it aims to improve the user's skills. Like OpenEssayist, Writing-Pal also uses interactive exercises to promote understanding. Writing-Pal is very different from OpenEssayist in terms of its underlying linguistic technologies and the design of its exercises. There is educational research that argues that using summaries in formative feedback on essays is very helpful for students. Nelson & Schunn, (2009) concluded that summaries make effective feedback because they are associated with understanding. They found that understanding of the problem concerning some aspect of an essay was the only significant mediator of feedback implementation, whereas understanding of the solution was not. Nelson & Schunn meant by the term 'summaries' the traditional notion of a short précis, and also some simpler representations, such as lists of key topics. Therefore automatic summarisation techniques became the main thrust of OpenEssayist development. An important consequence of this choice meant that OpenEssayist is domain independent. This also sets OpenEssayist apart from existing automated essay scoring systems.

Some of the existing technical systems that provide automated feedback on essays for summative assessment have been reviewed (Alden Rivers et al, 2014). Such systems focus on assessment rather than feedback, which is where OpenEssayist is different in providing hints for action for students to improve their essays. See Figure 1.

Fig 1: Screen shot of analysis of essay by OpenEssayist

In other research, hints have been used but have been given as responsive prompts, when students have requested help for a certain task or problem (e.g. Aleven et al., 2010), rather than as broad supportive information before starting tasks. In the study by Aleven and colleagues, the researchers focused on “help-seeking behaviour”, in considering when students requested the hints in order to gradually arrive at the answer, compared with those who were using hints to understand the question and how best to respond.

The purpose and design of OpenEssayist are very different from existing automated assessment systems. The system is primarily focused on user understanding and self-directed learning, rather than on essay improvement, and it engages the user on matters of content, rather than pointing out failings in grammar, style, and structure.

When the system was used with Open University postgraduate students, a wide variety of usage was found with respect to how long they engaged with the system and which features they accessed. Some continued to access the system and submit drafts after the course. Features concerning key words were popular, followed by highlighting key sentences and extracting these as a summary. A significant correlation was found between students' grades for Essay 1 and the number of drafts they submitted. Perhaps those students who submitted more drafts gained higher grades, or those students who tend to get higher grades also engaged more with the process of submitting drafts. We also found that this cohort of students, who had access to OpenEssayist, achieved significantly higher overall grades than the previous cohort, who did not. OpenEssayist has been used with students on a computer science course at Hertfordshire University in the UK and with students writing a research methods report at Dubai University. This is a learning analytics tool which has been rolled out in practice, and which has yielded evidence that students can and do benefit from using such a system in writing their academic assignments.

5. Open Comment

There is a recognition that e-assessment accompanied by an appropriate feedback to the student is beneficial for learning (DiBattista et al, 2004; Pitcher et al., 2002; Whitelock & Raw, 2003). Distance Learning too is forging ahead with electronic delivery of courses together with addressing the complexities of e-assessment for large cohorts of students. The Open Comment project sat within an external demand for electronic assessment from policy makers (see Whitelock & Brasher, 2006) and the Arts disciplines where analysis of more free text student responses were required rather than systems where students ticked multiple choice answers to questions. Open Comment, unlike OpenEssayist, was designed for a specific knowledge domain and cannot be used for any subject.

Free text response processing is at the cutting edge of linguistics. Certainly a completely human-like response to free text is still beyond the state-of-the art, but experience has shown that sometimes it is possible to provide effective responses based on surface features of a free text response. Carefully constructed language, conversational in form, can be even more important to guiding learning than the content being communicated (Holmberg, 1983). Instead of providing feedback on the answer, the project Open Comment's approach was, to some extent like ELIZA (Weizenbaum, 1966), to couch just enough analysis of the text in reflective language to help the learner assess their own work.

The specific objective of the project was to construct some simple tools in the form of Moodle extensions that allow a Moodle author to ask free-text response questions that can provide a degree of interactive formative feedback to students. In parallel with this was the aim to begin to develop a methodology for constructing such questions and their feedback effectively, together with techniques for constructing decision rules for giving feedback.

Open Comment provided a formative feedback technology designed to be integrated in the Moodle virtual learning environment. It delivers a simple system allowing questions to be written in Moodle, and for students' free text responses to these questions to be analysed and used to provide individually customised formative feedback. Open Comment is related to traditional free text assessment technologies, such as the ETS e-rater system and Landauer et al.'s (1998) IEA, although it has a very different emphasis. In particular, it makes no attempt to provide grading information; instead, it provides reflective feedback, designed to guide the students in their learning.

5.1 Pedagogical principles driving the feedback engine

This section reports on the pedagogical principles which drove Open Comment's development since the pedagogical rationale for this system was to engage students in a series of electronic formative assessment tasks that would provide more free text entry with automatic feedback. This would

promote a more challenging experience for the students than just checking their learning for revision purposes and promote a more personalised learning environment for self-reflection.

The guidance text arose from an analysis of what feedback actually was, and how learners used it. Throughout the development work, Whitelock & Watt (2008) worked closely with expert tutors in several Arts disciplines, using a range of techniques to elicit the processes they used to provide appropriate feedback. These ranged from role play (becoming a student) through to analysing collections of real answers and constructing sample solutions.

A preliminary analysis of 68 History assignments together with 100 plus assignments from different disciplines revealed a common pattern of tutor responses. These were clustered around the main categories of praise, advice on structure and presentation, particular misunderstandings, and developing and understanding particular issues.

The underlying model of feedback centred around:

- Identification of salient variables
- A description of these variables
- Identification of trends and relationships between these variables

The result of these analyses were formalised as an operational model for formative feedback generation, as set out in the Table 1 below.

Stages of Analysis by computer of students' free text entry for Open Comment	Advice with respect to Content	Socio-Emotional Support	Stylised Example
STAGE 1a: DETECT ERRORS E.g. Incorrect dates, facts. (Incorrect inferences and causality is dealt with below)	Instead of concentrating on X, think about Y in order to answer this question	Recognise effort (Dweck) and encourage to have another go	You have done well to start answering this question but perhaps you misunderstood it. Instead of thinking about X which did not..... Consider Y
STAGE 1b: IF NO INCORRECT STATEMENTS GO TO 2			A good start.....
STAGE 2a: REVEAL FIRST OMISSION	Consider the role of Z in your answer	Praise what is correct and point out what is missing	Good but now consider the role X plays in your answer
STAGE 2b: REVEAL SECOND OMISSION	Consider the role of P in your answer	Praise what is correct and point out what is missing	Yes but also consider P. Would it have produced the same result if P is neglected?
STAGE 3: REQUEST CLARIFICATION OF KEY POINT 1	Explain X more fully What do you mean by X	Confirm and concur about what is correct encourage to take the analysis further	
STAGE 4: REQUEST FURTHER ANALYSIS OF KEY POINT 1	Analyse X more fully	Confirm and concur about what is correct encourage to take the analysis further	Very interesting point – X is very complex perhaps it would have been effective to look at things slightly differently and consider how

(Stages 3 and 4 repeated with all the key points)			X contributes to Y
STAGE 5: REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS MISSING	Request the conclusion that can be drawn from X.	Praise effort and reiterate progress is being made	This is a sound description but it would be good if you explain what X is contributing to this situation.
STAGE 6: REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS NOT COMPLETE	What is X causing in this situation?	Reaffirm progress but encourage student to take the analysis process one step further	Yes what you have written is correct but can you elaborate and explain what X means?
STAGE 7: CHECK THE CAUSALITY	What is X, Y and Z causing in this situation?	Praise persistence and effort and ask the user to think about the reasoning behind a particular response.	You are certainly improving your answer to this question. Well done. In order to improve your answer further could you say something about the role X played in Y I'm thinking particularly of the following example where X was seen with respect to Z.
STAGE 7: REQUEST ALL THE CAUSAL FACTORS ARE WEIGHTED	Do X, Y and Z contribute in the same way to producing situation C, i.e. do the variables have equal weighting	Praise persistence and effort and ask the user to think about the importance and relative weightings of the causal factors	You have made a good stab at this question. From your answer I think you are allowing a considerable role to X. Does this mean you accept that X alone causes Y

Table 1: Operational feedback model for Open Comment

This model, as illustrated by Table 1, operates by and large through a sequential set of rules identifying sources of evidence within the student's response, and escalating in level of analysis, in some sense following Anderson, Krathwold, and Bloom's (2000) revised taxonomy of educational objectives. Importantly, also, there is a strong causal element to many of the rules. These rules were implemented in a bespoke feedback engine within Open Comment. An example of this feedback in the Open Comment system can be found in Figure 2.

Figure 2: Open Comment Arts

Upon closer inspection, Table 1 reveals specific advice with respect to the content of the student answer and also has a socio-emotional dimension, where the student effort is recognised and praise given for what has been correctly answered. This design approach was based upon the research of Mueller & Dweck (1998) who found that praising just ability can hinder the learner, but praising effort can motivate students to continue with their studies. This type of feedback can promote a growth mindset and lead to a lack of tension when learning, as the students know they can improve if they stretch themselves and confront obstacles as challenges (Dweck, 2008).

An important result from this project has been an increased understanding of the differences between even closely related disciplines. In both History and Philosophy, as with many humanities and social sciences, there is a greater emphasis on developing each student's ability to reason, and to use arguments and evidence in ways that are in keeping with a discipline-specific methodological ethos. Questions could rarely be taken at face value – especially in the more advanced levels. Open Comment feedback focused far more on evidence than on getting the answer right. In the future effective development of formative feedback technologies in these disciplines is totally dependent on effective involvement of tutors with both pedagogical and domain expertise.

6. Open Mentor

One of the challenges of today's education is that students are expecting better feedback, more frequently, and more quickly. Unfortunately, in today's educational climate, the resource pressures are higher, and tutor feedback is often produced under greater time pressure, and often later. Although feedback is considered essential to learning, what is it and how can tutors be supported to provide pertinent feedback to their students when automatic feedback is unavailable?

Human feedback is, put simply; additional tutoring that is tailored to the learner's current needs. In the simplest case, this means that there is a mismatch between students' and the tutors' conceptual models, and the feedback is reducing or correcting this mismatch, very much as feedback is used in cybernetic systems. This is not an accident, for the cybernetic analogy was based on Pask's (1976) work, which has been a strong influence on practice in this area (e.g., Laurillard, 1993).

One of the problems with tutor feedback to students is that a balanced combination of socio-emotive and cognitive support is required from the teaching staff, and the feedback needs to be relevant to the assigned grade. Is it possible to capitalise on technology to build training systems for tutors in Higher Education, that will support them with their feedback to students, and which will encourage their students to become more reflective learners? Since feedback is very much at the cutting edge of personal learning, this OpenMentor project set out to see how it could work with tutors to improve the quality of their feedback. To achieve this Open Mentor was developed as an open source tool which tutors can use to analyse, visualise, and compare their use of feedback.

With Open Mentor, feedback is not seen as error correction, but as part of the dialogue between student and tutor. This is important for several reasons: first, thinking of students as making errors is unhelpful – as Norman (1988) says, errors are better thought of as approximations to correct action. Thinking of the student as making mistakes may lead to a more negative perception of their behaviour than is appropriate. Secondly, learners actually need to test out the boundaries of their knowledge in a safe environment, where their predictions may not be correct, without expecting to be penalised for it. Finally, feedback does not really imply guidance (i.e., planning for the future) and OpenMentor has been designed to incorporate that type of support without resorting to the rather clunky ‘feed-forward’.

In order to provide feedback, Open Mentor has to analyse the tutor comments. So how could these comments be meaningfully grouped together? The classification system used in Open Mentor was based on that of Bales (1950). Bales’s system was originally devised to study social interaction, especially in collaborating teams; its strength is that it brings out the socio-emotive aspects of dialogue as well as the domain level. In previous work (Whitelock et al., 2004) found that the distribution of comments within these categories correlates very closely with the grade assigned.

Bales’s model provides four main categories of interaction: positive reactions, negative reactions, questions, and answers. These interactional categories illustrate the balance of socio-emotional comments that support the student. We found (Whitelock et al., 2004) that tutors use different types of questions in different ways, both to stimulate reflection, and to point out, in a supportive way, that there are problems with parts of an essay. These results showed that about half of Bales’s interaction categories strongly correlated with grade of assessment in different ways, while others were rarely used in feedback to learners. This evidence of systematic connections between different types of tutor comments and level of attainment in assessment was the platform for the current work.

Categories		Specific Examples
	Positive Reactions	
A1	1. Shows solidarity	Jokes, gives help, rewards others
A2	2. Shows tension	Laughs, shows satisfaction
A3	release	Understands, concurs, complies, passively accepts
	3. Shows agreement	
	Attempted Answers	
B1	4. Gives suggestion	Directs, proposes, controls
B2	5. Gives opinion	Evaluates, analyses, expresses feelings or wishes
B3	6. Gives information	Orients, repeats, clarifies, confirms
	Questions	
C1	7. Asks for information	Requests orientation, repetition, confirmation, clarification
C2	8. Asks for opinion	Requests evaluation, analysis, expression of feeling or wishes
C3	9. Asks for suggestion	Requests directions, proposals
	Negative Reactions	

D1	10. Shows disagreement	Passively rejects, resorts to formality, withholds help
D2		
D3	11. Shows tension	Asks for help, withdraws
	12. Shows antagonism	Deflates others, defends or asserts self

Table 2: Bales's Interaction categories

The advantage of the Bales model is that the classes used are domain-independent – we used this model to classify feedback in a range of different academic disciplines, and it has proven successful in all of them. An automatic classification system, therefore, can be used in all fields, without needing a new set of example comments and training for each different discipline.

Others (e.g., Brown & Glover, 2006) have looked at different classification systems, including Bales, and from these developed their own to bring out additional aspects of the tutor feedback, bringing back elements of the domain. In practice, no (useful) classification system can incorporate all comments. We selected, and still prefer, Bales because of its relative simplicity, its intuitive grasp by both students and tutors, and because it brings out the socio-emotive aspects of the dialogue, which is the one aspect tutors are often unaware of.

A second point is that Bales draws out a wider context: the team found that as they started to write tools that supported feedback, they began to question the notion of feedback itself. Instead, the concept seemed to divide naturally into two different aspects: learning support and learning guidance. Support encourages and motivates the learner, guidance shows them ways of dealing with particular problems.

6.1 Questions which tested the underlying pedagogical model for Open Mentor

Previous work by Whitelock, Watt, Raw & Moreale (2004) on student feedback has postulated that work that is awarded high grades should attract feedback from tutors that is high in praise, has few questions and does not ask the student to reflect on their work. Conversely, work that is awarded low grades should attract less praise, more questions and suggestions and invite more reflection. A number of questions in the Open Mentor Evaluation Study are able to throw light on these postulated outcomes and the results are summarised below.

A significant majority of both students and tutors respondents indicated that they expected high grades to attract more positive comments and low grades to attract more answers, suggestions and questions. Tutors gave a strong indication that they expected assessments with low grades to attract negative comments. Student responses followed a similar trend that was however not statistically significant. Students also indicated strongly that they expected no difference. All these findings support the pedagogical model postulated by Whitelock et al.

A further analysis, using cross tabulation revealed:

- Both students and tutors who feel that low grades would result in more questions also indicated that low grades would attract more answers
- Tutors who judged that high grades attract more positive comments also indicated strongly that low grades attract more answers and suggestions
- Tutors who felt that low grades attract more questions also indicated that low grades attract negative comments
- Both students and tutors felt that lower grades should attract more detailed comments and a deeper level of explanation. Higher grades should attract more positive comments

These findings from both groups of stakeholders supported a pedagogically driven development process which is described below.

OpenMentor was conceived as a tool to support tutors' feedback practices by classifying comments added to an assignment using Bales interaction analysis taxonomy (see Table 2) and reporting the results of the analysis in summarized views. Summary views show the proportion of the actual number of comments given by the tutors versus an ideal number. This calculated ideal is based on grade distribution and total comments included in the assignment, making the analysis unique to the student, tutor and feedback comments provided. Under Bales taxonomy, tutors' feedback comments are classified as Positive, Questions, Negative and Teaching Points. Examples of text identified by OpenMentor when classifying comments can be seen in Figures 3 and 4.

Figure 3: Screen shot of the analysis and categorisation of tutor comments by the Open Mentor system

Figure 4: Screen shot of Open Mentor comment analysis

OpenMentor has been used in anger by Southampton University and King's College London (Whitelock et al, 2012a; 2012b) and improvements made under the auspices of the OMTetra project. One of the important outcomes of the OMTetra project and the dissemination of OpenMentor is the positive effect in tutors' feedback practice, which would reflect on students' learning and performance. By supporting tutors' feedback practices through a strong formative function where the tutor can use the output of the system (reports and classifications) to engage in reflection about the quality and appropriateness of his/her feedback, students are more likely to receive feedback that is ultimately useful. Interestingly however, is the fact that students may also need to receive a form of training to interpret their tutors' feedback in order to benefit from receiving good quality feedback (Buhagiar, 2012). Further development of OpenMentor may include a student module where learners are asked to make notes on how they made use of their tutors' feedback. These notes could then be read by the tutor and mismatches between intended purpose of the feedback provided and that interpreted by the student are negotiated.

For tutors, there are significant opportunities in the use of OpenMentor as an academic development tool as it can generate dialogue about effective feedback between (a) tutors and academic developers and (b) peer reviewers during 'peer observation' of assessment practice. Consequently qualitative and quantitative outputs of the system which have been perceived as very useful during the pilots can be complemented by the function of the tool as generator of discussion and reflection on assessment practice. For students the tool can play a significant role in generating a dialogue between tutors and students about feedback and help them to close the loop (Sadler, 1989). This dialogue can achieve a consensus and a better understanding of standards of quality in student assessed work.

7. Conclusions

It has been proposed by Nelson & Schunn (2009) that there are 'mediators' that operate between the provision of feedback features, and implementation of suggestions. These authors addressed these mediators as 'understanding feedback' and 'agreement with feedback'. They suggested cognitive feedback factors are most likely to influence understanding, and affective factors are more likely to influence agreement. These are then said to influence implementation. Their results therefore showed a focus on how understanding feedback is critical to implementing suggestions from feedback. Thus it is important in designing course resources that we consider how to increase the likelihood that feedback is understood, if we want students to make use of it in current and future work – to learn from it (and improve performance) by *understanding* it, rather than just *improving one-off performance* by *blind implementation*. These proposals support the findings from the OpenMentor and Open Comment projects where socio-emotive support is recognised, together with cognitive assistance to provide students with "Advice for Action".

Equally important is the issue of students' 'mindsets', in their capacity for learning and improving performance, and in terms of students' beliefs that change is possible (Dweck, 2008; Yeager et al, 2013). The researchers refer to and contrast a 'fixed mindset', where students believe intelligence is relatively predetermined and finite; compared to a 'growth mindset', where students believe they can change their capacity and capabilities, respond to challenges, and try something again which they may initially find difficult. Thus students need to be given feedback that supports them in understanding requirements, but that also motivates them to believe they can make changes and improve their own performance. When such feedback can be given in a non-threatening way, for instance through live, personal use of an automated feedback system before formal submission, students may feel empowered that they can implement points raised in formative feedback to realize genuine improvements in performance. Both Open Comment and OpenMentor has taken on board the notion of changing mindsets in the feedback that is offered to both students and tutors alike.

The OMTetra project was successful in taking up Open Mentor and completing its transfer into two Higher Education Institutions. Interest shown by tutors from these institutions has translated into ideas to facilitate assignment analysis through Open Mentor and to encourage adoption of the system across institutional departments. Further development of Open Mentor features and promotion for adoption of the system at a larger scale are on-going efforts that will ensure that Open Mentor has an impact on a core task of HEIs: the delivery of quality feedback that will support the teaching and learning process.

OpenEssayist is unique in being a content-free tool that has been developed to offer automated feedback on students' draft essays, rather than an assessment on their finished work. OpenEssayist

is a system that offers opportunities for students to engage with and reflect on their work, in any subject domain, and to improve their work through understanding of the requirements of academic essay writing. In trial use of the system in a genuine Open University course, students made use of it to varying degrees (Whitelock et al, 2015), which is perhaps likely with any study resource. From Whitelock et al's analysis the team were also able to conclude that a significant positive correlation exists in this sample of students and the number of drafts submitted and the final grades for these essays. Another finding was that students who had access to OpenEssayist achieved significantly higher grades for this course than the previous year of students, who had no such access.

Furthermore OpenEssayist is a system that has been shown to offer opportunities for students to engage with and reflect on their work, in any subject domain, and to improve their work through understanding of the requirements of academic essay writing. In trialling use of the system in a genuine course, it was found that students made use of it to varying degrees, which is perhaps likely with any study resource. Those who took the time to explore system affordances and what they could be used for however tended to report more positively on its perceived value.

In today's educational climate, with the continued pressure on staff resources, making individual learning work is always going to be a challenge. But it is achievable, so long as we manage to maintain our empathy with the learner. Tools can help us achieve this by giving us frameworks where we can reflect on our social interaction, and ensure that it provides the emotional support as well as the conceptual guidance that our learners need.

8. Future Work

OpenEssayist has many potential advantages for students and tutors, which will benefit from further research and exploration. As OpenEssayist is designed to offer feedback to students during the drafting process, this has considerable implications for supporting students to improve their work, and also supporting students to believe that they can improve their academic work. This is no small feat for learners who may often feel isolated and stretched trying to squeeze study around other commitments and demands on their time.

Assessment with automatic feedback in Higher Education and any vocational training environment is a far more widespread issue than was fully realised by Whitelock et al (2015). After building OpenEssayist, it became apparent that there are many other potential applications for this technology. These include:

- Providing students with formative feedback on their assessments, with feedback properly adjusted to the students' needs
- Supporting the review process in academic conferences and competitive project proposals
- Automated generation of high quality reports (both in content and in presentation) based on complex data

With respect to OpenMentor, in the future, the current taxonomy used to analyse feedback in OpenMentor can be complemented with a dynamic algorithm that 'learns' from tutors feedback and classifies text using natural language processing techniques. This addition to the analysis algorithm of OpenMentor should address the needs of individual institutions where feedback practice is aligned to that of the culture of the organization. Our assumption, after the lessons learned from the implementation of OpenMentor in two Higher Education Institutions (Whitelock et al, 2012a; 2012b), is that the more configurable OpenMentor is, the more attractive it will be to disseminate its use across institutions.

Technology to enhance assessment and feedback is still developing but the problems are not technical: feedback coupled with assessment raises far wider social issues, and technologists have struggled in the past to resolve these issues with the respect they deserve. Automatic feedback is starting to deliver potential improvements; but there is still much work to be done.

Acknowledgements

The OpenEssayist Research was supported by the Engineering and Physical Sciences Research Council (EPSRC, grant numbers EP/J005959/1 & EP/J005231/1). Thanks are also due to the SAFeSEA team who produced OpenEssayist, namely John Richardson, Alison Twiner, Debora Field and Stephen Pulman. I would also like to thank Stuart Watt and the JISC for their support in the development of OpenMentor. Stuart Watt deserves special thanks as he also developed the Open Comment system.

References

- Alden, B., Van Labeke, N., Field, D., Pulman, S., Richardson, J. T. E., & Whitelock, D. (2014). Using student experience as a model for designing an automatic feedback system for short essays. *International Journal of e-Assessment*, 4(1), article no. 68.
- Alden Rivers, B., Whitelock, D., Richardson, J.T.E., Field, D., & Pulman, S. (2014). Functional, frustrating and full of potential: learners' experiences of a prototype for automated essay feedback. *Proceedings of Computer Assisted Assessment international conference: Research into eAssessment*, Zeist, the Netherlands, 30 June – 1 July.
- Aleven, V., Roll, I., McLaren, B.M., & Koedinger, K.R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist*, 45, (pp. 224-233). doi:10.1080/00461520.2010.517740
- Anderson, L.W., & Krathwohl, D.R. (eds.) (2000) *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon.
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. Paper presented at the 2nd International Conference on Learning Analytics and Knowledge, April 29th – May 2nd, Vancouver, BC, Canada. ACM 978-1-4503-1111-3/12/04.
- Bales, R.F (1950) A set of categories for the analysis of small group interaction. *American Sociological Review*, 15:257-63
- Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: A key to improving student learning? In *Studies in Higher Education*, 36, (pp. 671-687). doi:10.1080/03075071003731135
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bloom, B. S. (1971). *Handbook on formative and summative evaluation of student learning*. New York, NY: McGraw-Hill Book Company.
- Brown, E., & Glover, C. (2006). Written feedback for students: too much, too detailed or too incomprehensible to be effective? *Bioscience Education e-Journal*, 7(3).
- Buhagiar, M.A. (2012), Mathematics student teachers' views on tutor feedback during teaching practice, *European Journal of Teacher Education*, iFirst Article, 1-13
- Burstein, J. & Marcu, D. (2003). A machine learning approach for identification of thesis and conclusion statements in student essays. *Computers and the Humanities*, 37(4):455–467.
- Burstein, J., Chodorow, M. & Leacock, C. (2003). CriterionSM online essay evaluation: An application for automated evaluation of student essays. In J. Riedl and R. Hill, editors, *Proceedings of the Fifteenth Conference on Innovative Applications of Artificial Intelligence*, pages 3–10, Cambridge, MA.MIT Press.
- Chickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate

education. *American Association of Higher Education Bulletin*, 39(7), 3–7.

Dai, J., Raine, R.B., Roscoe, R.D., Cai, Z. & McNamara, D.S. (2011). The Writing-Pal tutoring system: Development and design. *Journal of Engineering and Computer Innovations*, 2(1):1–11. ISSN 2141-6508 2011 Academic Journals.

DiBattista, D., Mitterer, J.O. & Leanne, G. (2004) Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. *Teaching in Higher Education*, Vol. 9, No. 1, pp.17-28, ISSN 1356-2517

Dweck, C. (2008). *Mindset: The new psychology of success*. New York: Ballantine Books.

Evans, C. (2013). Making sense of assessment feedback in Higher Education. In *Review of Educational Research*, 83, (pp. 70-120). doi:10.3102/0034654312474350

Field, D., Pulman, S., Van Labeke, N., Whitelock, D. & Richardson, J.T.E. (2013). Did I really mean that? Applying automatic summarisation techniques to formative feedback. In: Angelova, G., Bontcheva, K. & Mitkov, R. (eds.) *Proceedings International Conference Recent Advances in Natural Language Processing*, 7-13 September, 2013, Hissar, Bulgaria. Association for Computational Linguistics, pp. 277-284.

Franzke, M. & Streeter, L.A. (2006). Building student summarization, writing and reading comprehension skills with guided practice and automated feedback. White paper, Pearson Knowledge Technologies.

Hattie, J., & Timperley, H. (2007). The power of feedback. In *Review of Educational Research*, 77, (pp. 81–112). doi:10.3102/003465430298487

Holmberg, B. (1983) Guided didactic conversation in distance education. In D. Sewart, D. Keegan, and B. Holmberg (Eds.), *Distance education: International perspectives* (114-122). London: Croom Helm.

Landauer, T.K., Laham, D. & Foltz, P.W. (2003). Automatic essay assessment. *Assessment in Education: Principles, Policy and Practice*, 10(3):295–308.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3), 259-284.

Landauer, T. K., Laham, D., & Foltz, P. W. (2003). Automated scoring and annotation of essays with the Intelligent Essay Assessor. *Automated essay scoring: A cross-disciplinary perspective*, 87-112.

Laurillard, D. (1993). *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*. London: Routledge.

Mayfield, E., & Rosé, C. P. (2013). LightSIDE: Open Source Machine Learning for Text. In M. D. Shermis & J. C. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 124-135). Oxon: Routledge.

McNamara, D.S., Raine, R., Roscoe, R., Crossley, S., Jackson, G.T., Dai, J., Cai, Z., Renner, A., Brandon, R., Weston, J., Dempsey, K., Lam, D., Sullivan, S., Kim, L., Rus, V., Floyd, R., McCarthy, P. & Graesser, A. (2011). The Writing- Pal: Natural language algorithms to support intelligent tutoring on writing strategies. In P.M. Mc-Carthy and Chutima Boonthum-Denecke, editors, *Applied Natural Language Processing and Content Analysis: Advances in Identification, Investigation and Resolution*, pages 298–311. IGI Global, Hershey, PA.

Merry, S., Price, M., Carless, D. & Taras, M. (2013) (eds.). *Reconceptualising Feedback in Higher Education: Developing Dialogue with Students*. London and New York: Routledge

Mueller, C. M. & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33-52.
<http://dx.doi.org/10.1037/0022-3514.75.1.33>

Narciss, S. (2013). Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. In *Digital Education Review*, 23, 7-26.

Nelson, M. M. & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instr. Sci.*, 37(4), 375-401.

Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. In *Studies in Higher Education*, 31, 199-218.
doi:10.1080/03075070600572090

Norman, D. (1988). *The psychology of everyday things*. New York: Basic Books.

Pask, G. (1976). *Conversation theory: applications in education and epistemology*. Amsterdam: Elsevier.

Pitcher, N., Goldfinch, J. & Beevers, C. (2002) Aspects of computer bases assessment in mathematics. *Active Learning in Higher Education*, 3(2),19-25.

Rivers, B. A., Whitelock, D., Richardson, J. T., Field, D., & Pulman, S. (2014). Functional, frustrating and full of potential: learners' experiences of a prototype for automated essay feedback. In: Kalz, Marco and Ras, Eric eds. *Computer Assisted Assessment: Research into E-Assessment. Communications in Computer and Information Science* (439). Cham, Switzerland: Springer, 40–52.

Rudner, L.M., Garcia, V. & Catherine Welch. (2006). An evaluation of the IntelliMetricSM essay scoring system. *The Journal of Technology, Learning, and Assessment*, 4(4).

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science* 18: 119-144.

Scott, D., Evans, C., Hughes, G., Burke, P. J., Watson, D., Walter, C. & Huttly, S. (2011). Facilitating transitions to masters-level learning: Improving formative assessment and feedback processes. Executive summary. Final extended report. London, UK: Institute of Education.

Simpson, O. (2012). *Supporting students for success in online and distance education*. Routledge, London, third edition.

Steffens, K. (2006). Self-regulated learning in Technology-Enhanced Learning Environments: Lessons from a European review. In *European Journal of Education*, 41, (pp. 353-380).
doi:10.1111/j.1465-3435.2006.00271.x

Taras, M. (2003). To feedback or not to feedback in student self-assessment. In *Assessment and Evaluation in Higher Education*, 28, (pp. 549-565). doi: 10.1080/0260293032000120415

Weizenbaum, J. (1966) ELIZA: A Computer Program for the Study of Natural Language Communication between man and machine. *Communications of the ACM*, 9(1), 36-45.

Whitelock, D. (2011). Activating assessment for learning: Are we on the way with Web 2.0. In M. J. W. Lee & C. McLoughlin (Eds.), *Web 2.0-Based-E-Learning: Applying Social Informatics for Tertiary Teaching* (Vol. 2, pp. 319-342): IGI Global.

Whitelock, D. (2015). Maximising Student Success with Automatic Formative Feedback for Both Teachers and Students. In: E. Ras and D. Joosten-ten Brinke (Eds.): *Computer Assisted Assessment. Research into E-Assessment*, Springer, pp. 142–148. ISBN 978-3-319-27703-5

Whitelock, D. & Raw, Y. (2003) Taking an electronic mathematics examination from home: what the students think. In C.P. Constantinou and Z.C. Zacharia (eds.) *Computer Based Learning in Science, New Technologies and their Applications in Education*, Vol. 1, Nicosia: Department of Educational Sciences, University of Cyprus, Cyprus, pp. 701-713, ISBN 9963-8525-1-3

Whitelock, D., Watt, S. N. K., Raw, Y., & Moreale, E. (2004). Analysing tutor feedback to students: first steps towards constructing an electronic monitoring system. *ALT-J*, 1(3), 31-42.

Whitelock, D. & Brasher, A. (2006) Developing a Roadmap for e-Assessment: Which Way Now? 10th International Computer Assisted Assessment Conference, Loughborough University, 4/5 July 2006, pp. 487-501. ISBN: 0-9539572-5-X

Whitelock, D. & Watt, S. (2008). Putting Pedagogy in the driving seat with Open Comment: an open source formative assessment feedback and guidance tool for History students. CAA Conference 2008, Loughborough University, 8/9 July 2008, Farzana Khandia (ed.) pp. 347-356.

Whitelock, D.M., Gilbert, L., Hatzipanagos, S., Watt, S., Zhang, P., Gillary, P. & Recio, A. (2012a). Addressing the Challenges of Assessment and Feedback in Higher Education: A collaborative effort across three UK Universities. In *Proceedings INTED 2012*, Valencia, Spain. ISBN: 978-84-615-5563-5

Whitelock, D., Gilbert, L., Hatzipanagos, S., Watt, S., Zhang, P., Gillary, P. Recio, A. (2012b). Assessment for Learning: Supporting Tutors with their Feedback using an electronic system that can be used across the Higher Education sector, In *Proceedings 10th International Conference on Computer Based Learning in Science, CBLIS 2012*, 26-29 June, Barcelona, Spain.

Whitelock, D., Field, D., Pulman, S., Richardson, J. T. & Van Labeke, N. (2014). Designing and testing visual representations of draft essays for higher education students. Paper presented at the 2nd International Workshop on Discourse-Centric Learning Analytics, 4th Conference on Learning Analytics and Knowledge, Indianapolis, India, USA.

Whitelock, D., Twiner, A., Richardson, J. T., Field, D., & Pulman, S. (2015). OpenEssayist: a supply and demand learning analytics tool for drafting academic essays. Paper presented at the Proceedings of the Fifth International Conference on Learning Analytics and Knowledge, Poughkeepsie, NY, USA.

Whitelock, D., Twiner, A., Richardson, J.T.E., Field, D. & Pulman, S. (2017). What types of essay feedback influence implementation: Structure Alone or Structure and Content? In: D. Joosten-ten-Brinke and M. Laanpere (Eds.): *Technology Enhanced Assessment TEA 2016*, CCIS 653, 1-16. DOI: 10.1007/978-3-319-57744-9_16

Yeager, D.S., Paunesku, D., Walton, G.M. & Dweck, C. (2013). How can we instill productive mindsets at scale? A review of the evidence and an initial R&D agenda. A White Paper prepared for the White House meeting on Excellence in Education: The Importance of Academic Mindsets