

OpenEssayist: A supply and demand learning analytics tool for drafting academic essays

Denise Whitelock
The Open University
Walton Hall, Milton Keynes
MK7 6AA, United Kingdom
+44 (0)1908 653777

denise.whitelock@open.ac.uk

Debora Field
Department of Computing
University of Oxford
Oxford, United Kingdom
debora.field.work@gmail.com

Alison Twiner
The Open University
Walton Hall, Milton Keynes
MK7 6AA, United Kingdom
+44 (0)1908 654811

alison.twiner@open.ac.uk

John T.E. Richardson
The Open University
Walton Hall, Milton Keynes
MK7 6AA, United Kingdom
+44 (0)1908 658014

john.t.e.richardson@open.ac.uk

Stephen Pulman
Department of Computing
University of Oxford
Oxford, United Kingdom
stephen.pulman@cs.ox.ac.uk

ABSTRACT

This paper focuses on the use of a natural language analytics engine to provide feedback to students when preparing an essay for summative assessment. OpenEssayist is a real-time learning analytics tool, which operates through the combination of a linguistic analysis engine that processes the text in the essay, and a web application that uses the output of the linguistic analysis engine to generate the feedback. We outline the system itself and present analysis of observed patterns of activity as a cohort of students engaged with the system for their module assignments. We report a significant positive correlation between the number of drafts submitted to the system and the grades awarded for the first assignment. We can also report that this cohort of students gained significantly higher overall grades than the students in the previous cohort, who had no access to OpenEssayist. As a system that is content free, OpenEssayist can be used to support students working in any domain that requires the writing of essays.

Categories and Subject Descriptors

K.3 [Computers and Education]: General

General Terms

Measurement, Performance, Design.

Keywords

Automated formative feedback, Educational performance, Academic essay writing, Online distance education, Natural language processing.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

LAK '15, March 16 - 20, 2015, Poughkeepsie, NY, USA
Copyright 2015 ACM 978-1-4503-3417-4/15/03...\$15.00
<http://dx.doi.org/10.1145/2723576.2723599>

1. INTRODUCTION

It has been suggested that learning analytics have the potential to be a ‘disruptive innovation’ [9] by empowering learners to have a greater understanding of their academic progress [1; 14]. Recent studies highlight the importance of continuous assessment for learning [15]. However, instructors are increasingly using open-essay type tasks to encourage critical thinking and reflection [11]. In contrast to multiple-choice testing, providing automatic feedback on open-essay type exercises is inherently difficult [8].

OpenEssayist is a real-time learning analytics tool that offers automated feedback for students drafting written essays [16]. It consists of a linguistic analysis engine that processes the text of an essay and a web application that uses the output of the analysis engine to generate feedback. The pedagogical challenge in the e-assessment of free text is how to provide meaningful ‘advice for action’ [15] to support students writing essays for summative assessment. In OpenEssayist, we have built a system that uses unsupervised graph-based ranking algorithms (following [10]) to extract key words, phrases and sentences from student essays. For this we use key phrase extraction and extractive summarization. In this paper we outline the system and describe observed patterns of activity as students engaged with the system for their assignments.

1.1 OpenEssayist

Many adults return to study after time in the workforce, and a significant period may have passed since their last experience of writing essays. It is not surprising that many find this task difficult and without adequate support may decide to leave. A system that can intervene and offer support between students’ draft and final essays could be valuable for students and tutors alike.

OpenEssayist processes open-text essays and offers feedback through key phrase extraction and extractive summarization. Each essay is automatically pre-processed using modules from the Natural Language Processing Toolkit [3]. These include several tokenizers, a lemmatizer, a part-of-speech tagger, and a list of stop words. Key phrase extraction identifies which phrases are most suggestive of the content, and extractive summarization identifies key sentences. These constitute the basis for providing feedback.

The system presents users with several kinds of feedback, including: identification of an essay’s most prominent words, with

graphical illustrations of their use across the essay; identification of the most representative sentences, with hints encouraging users to reflect on whether they express, in their view, the central ideas; and graphical illustrations of the essay's internal structure.

1.2 Feedback

It is generally agreed that feedback is central to learning [4]. It is however important to investigate how feedback can be offered in ways that support improvements in students' understanding of topics as well as their grades [6, 7]. Encouraging students to reflect on their work while they are engaging with the topic and task should have the most impact on students' understanding.

Feedback is most likely to influence performance positively when given during a task [5], rather than waiting until students submit a finished piece. In making use of feedback, it has been proposed that understanding problems raised within feedback about one's own work is a critical factor in implementing suggestions [12]. The researchers argued that understanding is more likely when those giving feedback locate problems, provide solutions, and summarize performance.

Feedback involves motivation, reinforcement and information [12]. 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' [12, p. 378].

It has been proposed that 'mediators' (especially 'understanding feedback' and 'agreement with feedback') operate between the provision of feedback features and implementation of suggestions [12]. 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.

Some of the existing technical systems that provide automated feedback on essays for summative assessment have been reviewed [2]. Such systems focus on assessment rather than feedback, which is where OpenEssayist is different. In considering the potential of OpenEssayist to offer the kind of support outlined above, in this paper we address the following research questions:

1. How did students use OpenEssayist, and which system features proved popular in terms of frequency of use?
2. What relationships are evident between system use and essay performance?
3. What rationales do students put forward for their use and non-use of particular system features?

2. METHOD

2.1 Course of Study

H817, 'Openness and innovation in elearning', is an optional module contributing to three postgraduate programs at the Open University. It was designed to introduce students to the latest

developments in educational communication and open learning, and so the adoption of OpenEssayist was highly appropriate.

2.2 Participants

During their assignment work in the 2013–14 academic year, 41 students enrolled on H817 accessed OpenEssayist at least once, using the login details they had been given. The system was opened to the students on January 27, 2014, with the final deadline that made use of this resource being May 5, 2014.

2.3 Procedure

Each student was asked to use OpenEssayist in writing their module essays within the specified timeframe. They were given individual logins and invited to explore the system, making use of the various features before submitting their essays. Submitted essays were marked by the usual team members. At the end of the module, students were invited to complete an online survey and to take part in a telephone interview about their use of OpenEssayist.

2.4 Data Analysis

Frequency analysis was conducted on the activity logs generated by students' use of OpenEssayist. From this we could view patterns of use over time and across the various system features. Further frequency analysis was carried out on the students' survey responses, and their written comments were reviewed. This was combined with data from an interview with one student after the module work was completed. Correlation analysis was performed across students' essay grades and data on system usage.

3. RESULTS

3.1 Draft Essays Submitted

Of the users for whom we have data on both when they accessed the system and how many drafts they submitted per session (27 of the 41 users), most (23 users) did not submit a draft on their first visit. The majority of users accessed the system on two, three or four different sessions (11 users, 8 users and 9 users respectively). The access times clustered around assignment submission dates, showing that OpenEssayist was used to support essay drafting.

3.2 Duration of Use

For 30 users, we have data on the duration of each session in which they were logged onto the system. From this we can calculate the total time they spent on OpenEssayist (although it is not possible to ascertain whether they were always engaged and 'active' in the system – they may have visited other websites or moved away from their computer for a period of time whilst still logged on), and the mean time spent per session (see Table 1).

Table 1. Mean session length

Mean session length	Number of users
Less than 10 minutes	18
10-30 minutes	6
31-60 minutes	1
61 minutes-2 hours	3
Over 2 hours	2

Here is an example of one student's use of the system. This user accessed the system on 10 occasions for a total time period of 31 minutes – giving a mean session time of 3 minutes. Within the 10 sessions they submitted seven drafts. Two drafts were submitted in a session lasting 10 minutes, one was submitted in a session lasting 8 minutes, and the remaining four drafts were submitted during sessions lasting no more than 4 minutes. This user was therefore very active in the system during these short periods – but were they using the time to draft, submit and revise lots of essays (with revision being done outside OpenEssayist), or were they submitting many drafts without making use of the system features? Clearly we can only speculate on this issue.

Although this is just an example, there is a complicated picture of usage. Some students used OpenEssayist evenly and sparsely, with many short sessions spread over several weeks. Others used the system in fewer but more concentrated chunks. Some used OpenEssayist very little overall, while others used the system – or at least were logged onto it – for long stretches of time.

3.3 System Features

Data was recorded from 35 users on how many times they accessed different system features. OpenEssayist is made up of three major components or 'landing screens' – 'essay', 'analysis' and 'graphics' (see Figure 1). Users are initially brought to these landing screens when they log in, from which they can choose different options. The first main component, 'essay', offers representations of students' texts concerning the essential elements of an essay: key words and sentences. The screens present to students the introduction, discussion and conclusion sections of their essay and identify where key words and sentences occur throughout the text. The second component, 'analysis', pulls the key words and sentences out of the essay text and offers students the opportunity to organize key words. These options aim to encourage students to consider whether their essay contains the

key concepts and development of argument that they intended. The third main component, 'graphics', includes two different visual representations of key words: as a word cloud, a very common technique for visual analysis of documents which gives prominence to frequent words [13]; and as a dispersion plot, showing the distribution of key words through the essay. The graphics landing screen offers two ways to view the word count in terms of how the word count is divided across the essay sections and how the word count aligns with the word limit.

All users submitting a draft accessed the essay landing screen, as this is where a user is automatically routed once they have uploaded a draft. All but one user accessed either the analysis screen or the graphics screen. Their use ranged from accessing one feature (one user), to accessing them all (seven users accessed all 10 features). The majority of users accessed seven or more features (23 users, see Table 2).

Of the 35 users and 10 features, this gives 350 potential opportunities for each student using each feature. Of these, 105 had the value of zero (no user accessed the feature), and a further 113 had the value of 1, indicating that a user accessed the feature but did not return to it. This leaves 132 instances in which users accessed a feature on two or more occasions. Indeed there were 21 instances of a user returning to a feature five or more times, up to a maximum of 10 visits (for 'highlight key words' and 'show word limit comparison').

Of the 11 users who accessed five or fewer features, most accessed these features only once (this made up 81% of access for this category), with the remainder returning to a feature 2–4 times (19%). Of the users who accessed more than five features, most accessed them 2–4 times (50% of access for this category), a further 39% of features were accessed once by these users, and the final 11% of features were accessed five or more times. So users who accessed more features tended to return to more features.

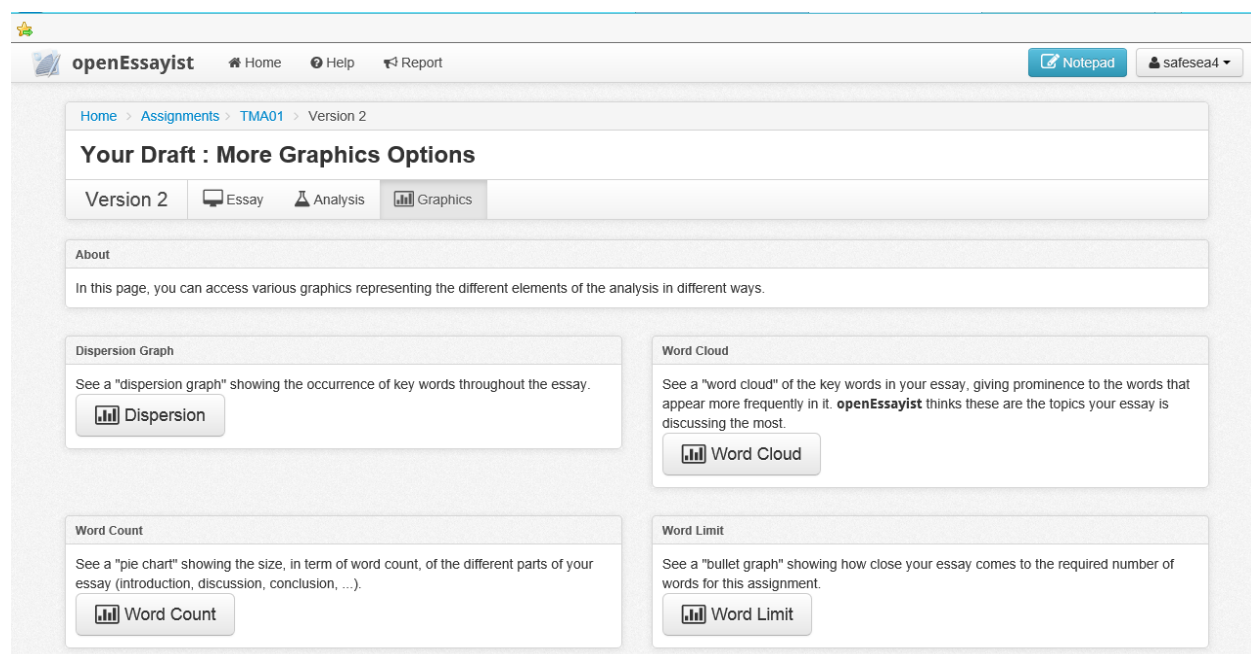


Figure 1 – The OpenEssayist 'graphics' landing screen, with 'essay' and 'analysis' tabs visible

Of the essay features, users preferred to show key words highlighted on their text, followed by highlighting embedding key sentences, rather than showing both of these aspects highlighted simultaneously on their text. Of the analysis features, users mostly requested a list of key words. Some also looked at the list of key sentences, which offered them a summary of their essay in text or importance order. Of the graphics features, the most popular was viewing the word cloud, closely followed by the dispersion graph.

In short, students largely took advantage of the key word and sentence options. These features highlight key concepts within an essay, and the representations allow students to consider whether its structure and content present a coherent argument. Students are given information on the spread of ideas through the essay and the connectedness and development of concepts across their introduction, discussion and conclusion.

3.4 Students' Reflections on OpenEssayist

Twenty students completed the survey regarding their use of OpenEssayist for H817. All 20 respondents indicated that they had used OpenEssayist for Essay 1, and 13 also reported using it for Essay 2. The majority of these students (16 or more) indicated that they had made use of the essay features. In terms of the key words, phrases and sentences views, comments from students who found these features useful included that they supported them in checking essay structure, checking their reference back to the essay question, and identification of repetition, as well as determining whether any elements they considered key might need to be emphasized more strongly if they had not appeared.

In terms of the graphics features, such as the word cloud and key word dispersion plot, 12 respondents indicated that they had used these views. Those who had not found these features helpful commented that they merely replicated information they had gained from other views or that they found them confusing. Those who appreciated these screens again suggested that they helped them to consider the structure of their essay visually, and to reflect on whether the represented essay matched their intentions.

3.5 Grades and Use of OpenEssayist

Grades were available for 30 students on Essay 1 and for 14 participants on Essay 2. Data from three participants was removed, as the time for which they were logged into the system was substantially longer than the other participants, implying that they were doing other things whilst still logged into the system.

Correlational analysis yielded three significant relationships. First, there was a positive relationship ($r = +0.41$) between grades for Essay 1 and number of drafts. This could mean that submitting more drafts leads to higher grades, or that brighter students submit more drafts. Second, there was a positive relationship ($r = +0.65$) between number of visits and number of drafts. Third, there was a positive relationship ($r = +0.60$) between mean time and total time. The relationship between number of visits and grades for Essay 2 was also significant by a one-tailed test ($r = +0.50$).

There was no significant difference found between the mean grade for Essay 1 and the mean grade for Essay 2. For the 13 students who submitted both essays, their mean grades were 70.9 for Essay 1 and 73.3 for Essay 2. We did however find a significant difference between the mean grade awarded for the module overall for this cohort of students (64.2), who had the opportunity to use OpenEssayist in preparing their assignments, and that of students in the previous cohort, who did not (53.7) ($p = .04$).

3.6 Interview Comments

We interviewed one student who had used OpenEssayist on this course. The student had used the system for both assignments and had used it for a further two since then. His first expectation was that it would give him initial feedback in the form of a grade before submission, so it was not initially what he had expected. In light of this he felt it was not as helpful to him as he hoped on his first essay. He continued to explore it, however, and realized the value of some of the features, and found it very helpful in cleaning up structure (in particular, the use of key phrases and sentences). He felt it had altered how he went about essay writing,

Table 2. Access to system features

	Essay			Analysis			Graphics			
	Highlight key words	Highlight key sentences	Highlight key words and sentences	Show extracted key words	Show extracted key sentences	Organize key words	Show key word cloud	Show key word dispersion plot	Show word count structure chart	Show word limit comparison
Users accessing this feature	28	29	23	30	28	15	24	28	19	21
Total accesses for this feature	115	99	49	105	66	26	103	96	47	59

in structuring and dividing the essay among the different sections as well as dividing his time across the various elements of the essay. He thought that his grades had improved, too.

4. DISCUSSION

There was wide variety in how frequently students used the OpenEssayist system, for how long, and which features they accessed. Some continued to access the system and submit drafts after the course. The majority accessed at least seven features, and those who accessed more features tended to return to more features. 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. CONCLUSIONS AND IMPLICATIONS

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 trialing use of the system in a genuine course, we 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. Use of a system such as OpenEssayist has many potential advantages that will benefit from further research. In particular, it can support students both to improve their academic work and also to *believe* that they can improve their work.

6. ACKNOWLEDGMENTS

This work is supported by the Engineering and Physical Sciences Research Council (EPSRC, grant numbers EP/J005959/1 & EP/J005231/1).

7. REFERENCES

- [1] Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-Gonzalez, M. Á., and Hernandez-Garcia, Á. 2014. Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Comput. Hum. Behav.*, 31, 542-550.
- [2] Alden Rivers, B., Whitelock, D., Richardson, J. T. E., Field, D., and 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 e-assessment*, Zeist, The Netherlands, June 30–July 1.
- [3] Bird, S., Klein, E., and Loper, E. 2009. *Natural language processing with Python*. Cambridge, MA: O'Reilly.
- [4] Black, P. and Wiliam, D. 1998. Assessment and classroom learning, *Assessment in Education*, 5(1), 7–74.
- [5] Butler, D. L. and Winne, P. H. 1995. Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.*, 65(3), 245–281.
- [6] Chickering, A. W. and Gamson, Z. F. 1987. Seven principles for good practice in undergraduate education. *American Association of Higher Education Bulletin*, 39(7), 3–7. <http://www.aahea.org/aahea/articles/sevenprinciples1987.htm> Accessed October 8, 2014
- [7] Kluger, A. N. and DeNisi, A. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.*, 119(2), 254-284.
- [8] Koedinger, K., Booth, J. L., and Klahr, D. 2013. Instructional complexity and the science to constrain it. *Science*, 342, 935-937.
- [9] Law, N. 2014. Is learning analytics a disruptive innovation? Paper presented at the *4th International Conference on Learning Analytics and Knowledge*, March 24–28, Indianapolis, USA.
- [10] Mihalcea, R. and Tarau, P. 2004. TextRank: Bringing order into text. *Proceedings of Empirical Methods in Natural Language Processing EMNLP* (pp. 404-411) <http://aclweb.org/anthology/W/W04/#3200> Accessed 8th October 2014
- [11] Narciss, S. 2013. Designing and evaluating tutoring feedback strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. *Digital Education Review*, 23, 7-26 <http://greav.ub.edu/der> Accessed July 17, 2014
- [12] Nelson, M. M. and Schunn, C. D. 2009. The nature of feedback: How different types of peer feedback affect writing performance. *Instr. Sci.*, 37(4), 375-401.
- [13] Paulovich, F. V., Toledo, F. M. B., Telles, G. P., Minghim, R., and Nonato, L. G. 2012. Semantic wordification of document collections. *Comput. Graph Forum*, 31(3), 1145-1153.
- [14] Tempelaar, D. T., Rienties, B., and Giesbers, B. in press. In search for the most informative data for feedback generation: Learning Analytics in a data-rich context. *Comput. Hum. Behav.*
- [15] Whitelock, D. 2011. Activating assessment for learning: Are we on the way with Web 2.0? In M.J.W. Lee and C. McLoughlin (Eds.), *Web 2.0-Based-E-Learning: Applying social informatics for tertiary teaching* (pp. 319-342). IGI Global.
- [16] Whitelock, D., Field, D., Pulman, S., Richardson, J. T. E., and Van Labeke, N. 2013. OpenEssayist: An automated feedback system that supports university students as they write summative essays. *The 1st International Conference on Open Learning: Role, Challenges and Aspirations*, The Arab Open University, Kuwait, November 25–27, 2013.