ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 2, Issue 2, pp: (271-277), Month: April-June 2014, Available at: www.researchpublish.com

Implementation of Automatic Taxanomy Tools to Improve Research Practice

A.Padmini[1] and Dr.R.Indra Gandhi[2]

[1]Research Scholar, Dept of Computer Applications, GKM College of Engineering and Technology, Chennai-63
[2]Professor, Dept of Computer Applications, GKM College of Engineering and Technology, Chennai-63

Abstract: **Academic writing, individual or collaborative, is an essential skill for today's graduates. Unfortunately, managing writing activities and providing feedback to students is very labour intensive and academics often opt out of including such learning experiences in their teaching. We describe the architecture for a new collaborative writing support environment used to embed such collaborative learning activities in engineering courses. In this paper provides tools for managing collaborative and individual writing assignments in large cohorts. It outsources the writing tools and the storage of student content to third party cloud-computing vendors (i.e. Google). We further describe how using machine learning and NLP techniques, the architecture provides automated feedback, automatic question generation and process analysis features.**

Keywords: **Collaborative learning tools, Homework support systems, Intelligent tutoring systems, Peer reviewing**

1. INTRODUCTION

Writing is important in all knowledge-intensive professions. Engineers, for example, spend between 20% to 40% of their workday writing, a figure that increases with responsibility of the position. It is often the case that much of the writing is done collaboratively. For example, Ede and Lundsford showed that 85% of the documents produced in offices and universities had at least two authors. This results is similar to those in others studies. Collaboration and writing skills are so important that accreditation boards such as the Accreditation Board in Engineering and Technology (ABET) require evidence that graduates have the "ability to communicate effectively". However, motivating and helping students to learn to write effectively before they graduate, particularly in collaborative scenarios, poses many challenges, many of which can be overcome by technical means. E. C. Thiesmeyer and J. E. Thiesmeyer etal. expressed an over the last 20 years, researchers within universities have been developing technologies feedback in academic writing and for enabling collaborative writing (CW), but work combining both automated feedback and CW has been scant.

2. RELATED WORKS

2.1 Collaborative writing

P. B. Lowry, A. Curtis etal. says that the features of computer-supported CW has received attention since computers have been used for word processing . two areas of research are particularly relevant for our project: research that analyses CW in terms of group work processes, focusing on issues such as process loss, productivity, and quality of the outcomeS; and research that studies CW in terms of group learning processes, focusing on topics such as establishing common ground, knowledge building, and learning outcomes. M. Scardamalia and C. Bereiter proposed in the second line of research (computer-supported collaborative learning, CSCL), writing is seen as a means to deepen students' engagement with ideas and the literature and for knowledge building by jointly developing a text or hypertext. In CSCL, in addition to knowledge building in asynchronous collaborative development of argumentative structures and texts has received much attention .

Because of the complexity of the CW process, explicit support needs to be provided, in particular for novice writers. Such support generally falls into one of three classes: specialized writing and document management tools, document analysis software, and team process support.

2.2 Writing for learning

M. Scardamalia, C. Bereiter etal. described writing for Learning (WfL), with variations such as Writing Across the Curriculum and Knowledge Building pedagogy , has attracted the interest of teachers and of researchers for more than thirty years. It has, for instance, seen widespread use in science education. We are proposing WfL not only because research has 'shown that it works' (although the empirical findings are, as usual, mixed but in particular because it can be flexibly employed in formal as well as non-formal learning settings. V. Prain and J. P. Gee proposed Futhermore, writing researchers have theorized and studied the intricate relations between congnition, interest, and identity in a holistic fashion before which makes it particularly relevant for engineering education. A number of reasons have been identified to explain why writing is an important tool for learning. J. R. Hayes and L. S. Flower engaged cognitive psychologists make the general argument that writing requires the coordination of multiple perspectives (content and audience) and the linearization thought, which might not be linear. For subject matter learning, this means that writing requires deep cognitive engagement with content, which will lead to better learning. J. P. Gee says from a discourse theory perspective, it has been argued that students must learn to understand and reproduce a professional community's traditional written discourse if they are to become members of that community. V. Prain explained pedagogy- based arguments for the value of writing is an important medium for reflection and, in the context of higher education, also a medium for developing epistemic orientations.

2.3 Autsomated essay feedback and scoring systems

Automated feedback systems have been studied for over a decade and most of these systems focus on individual writing, not on collaborative activities. Over this period techniques of Natural Language Processing (NLP) and Machine Learning have progressed substantially and automated writing tutors have improved simultaneously. M. Warschauer and P. Ware despite this progress, the value of automated feedback and essay scoring remains contested [26]. The increasing use of automatic essay scoring (AES) in particular by many institutions has created robust debates about accuracy and pedagogical value. Two recent books about accuracy and pedagogical value. M. D. Shermis, J. Burstein says two recent books discuss advance in AES, one taking a very supportive approach [27] and one providing a more critical debate [28]. R. A. Calvo and R. A. Ellis explained about Glosser [29] is an automatic feedback tool used within iWrite for selected subjects. It was designed to help students review a document and reflect their writing [29]. J. Villalon, P. Kearney, etal. said glosser uses textual data mining and computational linguistics algorithms to quantify features of the text, and produce feedback for the student [30]. This feedback is in the form of descriptive information about different aspects of the document. For example, by analysing the words contained in each paragraph, it can measure how thematically 'close' two adjoining paragraphs are. If the paragraphs are too 'far' this can be a sign of a lack of flow, and Glosser flags a small warning sign. As a form of feedback Glosser provides trigger questions and visual representations of the document. M. Warschauer and P. Ware defined other researchers have used techniques similar to those used in Glosser for Automatic Essay Assessment for building writing support tools. Criterion (by ETS Technologies), MyAccess (by Vantage Learning) and WriteToLearn by Pearson Knowledge Technologies are all commercial products increasingly used in classrooms [26]. These programs provide an editing tool with grammer, spelling, law-level mechanical feedback. They also provide resources such as thesaurus and graphic features, many of which would be available in tools such as MS Word. To our knowledge these tools do not have collaborative writing or process oriented support.

As far as we know, all the system reported in the literature are designed as stand-alone activities, normally used outside the context of a real class scenario. R. Ellis and R. A. Calvo, etal. analysed that would likely affect the conceptions that students have about the activity, and therefore the way they engage in it. Evidence shows that in collaborative and in writing activities [32, 33] this significantly affects the learning outcomes. J. B. Biggs expressed systems like iWrite, which afford collaborative writing activities that are embedded and constructively aligned [34] with the assessment and the learning outcomes, are more likely to be successful.

3. PROPOSED ARCHITECTURE

The 'iWrite' website provides student with information about their writing and submitting their assignments, and a complete solution for scaffolding the write-review-feedback cycle of a writing activity. Figure 1 shows its three sections, two of which (For Students' and 'For Academics') consist of content and interactive tutorials on developing students' understanding of different concepts and genres of writing. These consist of discipline specific tutorial exercise where students are introduced to writing concepts through examples written by others. Only the 'Assignments' section –which

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 2, Issue 2, pp: (271-277), Month: April-June 2014, Available at: www.researchpublish.com

supports to contextualize these writing concept in their own compositions – will be discussed here in detail. Service Level Agreement (SLA) ensures that assignment documents are always available. The architecture of the system is illustrated in Figure 2. The writing tools and activities, on the left hand side, are implemented with Google Docs, a cloud based office suite for editing documents, presentations and spreadsheets. The API provides programmatic access to the documents. The right hand side of Figure 2 shows the Assignment Manager, Glosser and WriteProc. Assignment Manager deals with the administration and scheduling of courses and writing activities.

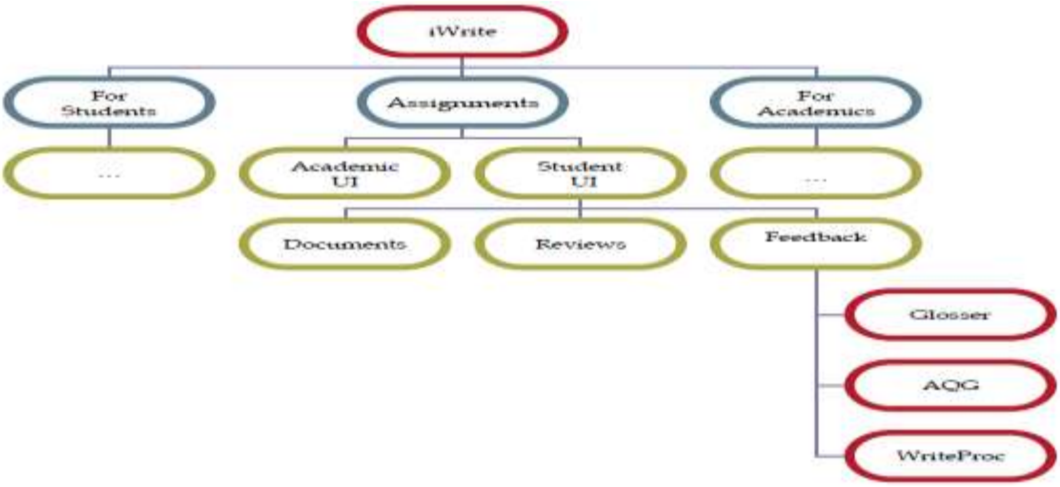

Figure 1: The iWrite information structure diagram

In addition, tools that analyse the document using NLP techniques provide additional functionalities. Automatic Question Generation (AQG) generates questions from templates based on the references used in a document. WriteProc is a tool for analysing students usage of Write in combination with the methodological process of their writing. Through the 'Assignments' tab of the website, students have access to the documents, feedback from instructors, peers and system, information about deadlines and so forth. These are shown in the top two boxes of the screen-shot of Figure 3. If the user is identified as an instructor for a particular course, additional features are provided (e.g. downloading a zip file with all the submitted assignments) as shown in the lower box of Figure 3.Both Glosser and WriteProc use TML, a multipurpose text mining library that implements the NLP and machine learning techniques that analyse actual content of the document revisions. TML provides a comprehensive set of text mining algorithms and scaffolds every stage of the text mining process. TML integrates the open source Apache Lucene search engine, the Stanford NLP parser and the Weka machine learning libraries, and is itself open source. TML provides functionalities for the pre-processing of documents, tokenising, stemming and stop-word removal. It maintains three corpora, adding each new document, at the sentence, paragraph and document level. In order the lag-time, all these are stored in a repository, along with the results of the text mining operations.

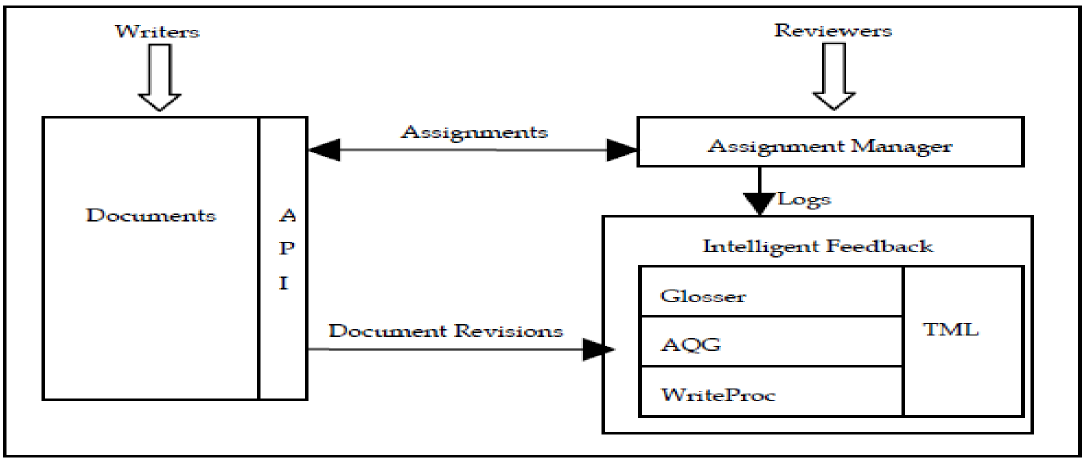

Figure 2: The iWrite architecture diagram

3.1 Assignment Manager

The Assignment Manager is designed to use cloud computing applications and their APIs. This means that the writing tool and the documents themselves are managed by a third party. This significantly reduces of the cost managing a system with large number of students, and a Assignment Manager handles all aspects of the assignment submission, peer-reviewing and assessment process. It uses the API provided to a Google Apps for Education account to administer user accounts and to create, share and export documents. The APIs operate using an Atom feed to download data over HTTP. Although the Assignment Manager is currently only integrated with the Docs service, there is the potential to incorporate other Google services into activities, such as Sites or Calendar. An abstraction layer also allows systems from other vendors to be added. Assignment Manager is administered through a web application based on Google Web Toolkit which facilities the creation of courses and writing activities. Each course has a list of students (and their contact information), maintained in a Google Docs spreadsheet and synchronized with Assignment Manager on request. Keeping this information in a spreadsheet allows course managers to easily modify enrolment details in bulk, and assign student to groups and tutorials. Assignment Manager maintains a simple folder structure of courses and writing activities on Google Docs. The permission structure of the folder tree is such that lecturers are given permission to view all documents in the course and tutors are given permission to view all documents of the students enrolled in their respective tutorials. A writing activities can specify a document type (i.e document, presentation, or spreadsheet), a final deadline along with optional draft and review deadlines, along with various other settings. A final copy of each submitted document is downloaded in PDF format and distributed to tutors and lecturers for marking.

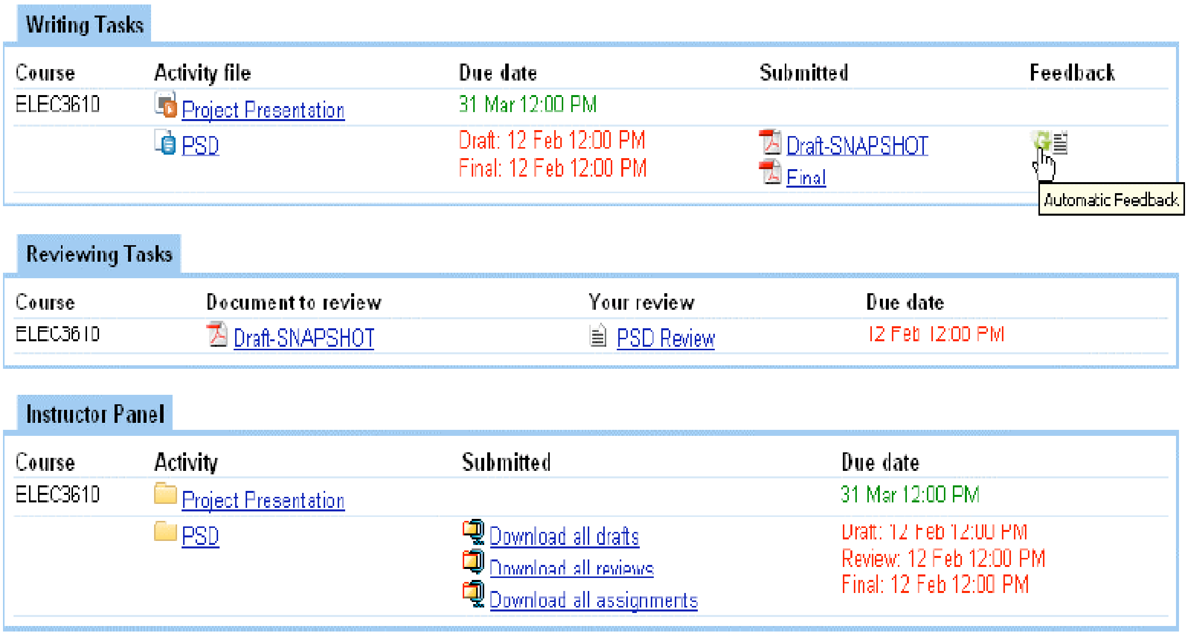

Figure 3: A screenshot of the Assignment Manager

3.2 Intelligent Feedback: automatic feedback, questions and process analysis

J. Villalon, P. Kearney, etal. prepared using the APIs the system has access to the revisions of any document. This allows new functionalities such as automatic plagiarism detection, automatic feedback, and automatic scoring systems to be integrated seamlessly with the appropriate version of the document. iWrite currently implements 3 such intelligent feedback tools, Glosser, AQG and WriteProc, to generate automatic feedback, questions and process analysis, respectively.

3.3 Glosser: Automatic Feedback Tool

J. Villalon, P. Kearney, etal. expressed Glosser is intended to facilitate the review of academic writing by providing feedback on the textual features of a document, such as coherence. The design of Glosser provides a framework for scaffolding feedback through the use of text mining techniques to identify relevant features for analysis in combination with a set of trigger questions to help prompt reflection. The framework provides an extensible plugin architecture, which allows for new types of feedback tools to be easily developed and put into production.

3.4 Automatic Question Generation (AQG)

M. Liu, R. A. Calvo, proposed the iWrite architecture includes a novel Automatic Question Generation (AQG) tool [36] that extracts citations from student's compositions, together with key content elements. For example, if the students use the APA citations style, author and year are extracted. Then the citations are classified using a rule-based approach.

3.5 WriteProc: Process Mining Tool

The autosave function in Google Docs acts as a version tracking functionality, saving documents every 30 seconds or so (as long as the student has written something in that period). This means that for each single document written by a student or team thousands of revisions are stored.

4. EVALUATION

We present here three different evaluation aspects. First we show a traffic analysis of iWrite, which is key to understand how to understanding how the tool is used and the writing Process involved. Second we analysed further how high achieving students differed from other students with respect to the way they worked on their collaborative writing assignment. Lastly we include some user feedback.

4.1 Writing process

The writing processes followed by students in different activities (or subjects) reflect their understanding of what is expected in the activity, their motivation and other educational factors. Often students' behaviour during an activity can be different from what instructors expect, and this variation may raise issues on how then activity is designed.

4.2 Relating aspects of iWrite use with student performance

We categorized students of the ENGG 1803 course into three groups, according to the mark they received for their collaborative writing assignment. We considered a low mark to be below one standard deviation from the mean mark, a medium mark to be within one standard deviation from the mean, and a high mark above the mean plus one standard deviation. Table 1 summarizes the details of these three groups.

TABLE 1

GROUPING OF ENGG1803 STUDENTS ACCORDING TO THEIR COLLABORATIVE WRITING ASSIGNMENT GRADES

Assignment Mark	N	Mean	Std. Dev.	Min	Max
Low	28	58.30	3.96	47.5	62.5
Medium	103	72.05	4.57	63.3	79.5
High	23	83.66	2.49	80.5	88.3
Total	154	71.28	8.47	47.5	88.3

We then compared these three groups in relation to the iWrite usage variables defined in the previous section, using ANOVA. We found that the four variables which were significantly related to grades were:userRevisions ($F=3.146$, $p=0.049$), teamRevisions ($F=3.388$, $p=0.025$), sessionsWriting ($F=6.381$, $p=0.003$), daysWriting ($F=7.948$, $p=0.001$). Posthoc analysis (using Tukey's HSD) revealed the following:

Students with low grades did more individual and group revisions compared to these with medium grades ($p=0.025$, and $p=0.026$, respectively)

Students with low and medium grades engaged in fewer writing sessions compared to students with high grades ($p=0.036$ and $p=0.001$, respectively)

Students with low and medium grades engaged in fewer writing days compared to students with high grades ($p=0.003$ and $p<0.001$, respectively).

These results indicate that it is not whether, but how, students used iWrite which made a difference in their CW performance.

4.3 User feedback

We collected informal feedback from the course lecturers who used iWrite. They were extremely positive about the experience. One course manager commented "An online assignment submission system will save us a lot of time sorting and distributing assignments. In addition, we send copies of a portion of our assessments to the Learning Centre, so online submission really minimises our paper usage".

5. CONCLUSIONS

The architecture for iWrite, a CSCL system for supporting academic writing skills has been described. The system provides features for managing assignments, group and peer-reviewing activities. It also provides the infrastructure for automatic mirroring feedback including different forms of documents visualization, group activity and automatic generation. We described aspects of its use with large cohorts, and comments from students and administrators. Whilst an evaluation of the system's impact on learning and the students' perceptions of writing are outside the scope of this paper, we analysed student use of iWrite in relation to student performance and found that the best predictors for high performance are the way students use iWrite, not necessarily whether they used the tool. This is an important finding that gives clear design guidelines for teachers as well as explicit good writing practices for students. Our future evaluation work will include showing this type of statistical information to instructors and inquire if the values are what they expected and how these data can be used to inform their pedagogical design.

REFERENCES

[1] M. L. Kreth, "A survey of the co-op writing experience of recent engineering graduates," IEEE Trans. Prof. Commun., vol. 43, pp. 137 – 152, June 2000.

[2] L. S. Ede and A. A. Lunsford, Singular texts/plural authors: Perspectives on collaborative writing: Southern lllinois Univ Pr, 1992.

[3] E. C. Thiesmeyer and J. E. Thiesmeyer, Editor. New York: Modern Language Association, 1990.

[4] T. J. Beals, "Between Teachers and Computers: Does Text-Checking Software Really Improve Student Writing?," English Journal, National Council of Teachers of English, vol. 87, pp. 67—72, 1998.

[5] P.Wiemer-Hastings and A. C. Graesser, "Select-a-Kibitzer: A computer tool that gives meaningful feedback on student composition, " Interactive Learning Environments, vol. 8, pp. Taylor\&Francis—169, 2000.

[6] D. Wade-Stein and E. Kintsch, "Summary Street: Interactive computer support for writing," Cognition and Instruction, vol. 22, pp. 333—362, 2004.

[7] P. B. Lowry, A. Curtis, and M. R. Lowry, "Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice., " Journal of Business Communication, vol. 41, pp. 66-99, 2004.

[8] G. Erkens, J. Jaspers, M. Prangsma, and G. Kanselaar, " Coordination processes in computer in computer supported Collaborative writing," Computers in Human Behavior, vol. 21, pp. 463-486, 2005.

[9] M. Scardamalia and C. Bereiter, "Higher Levels of Agency for Children in knowledge building: A Challenge for the Design of New Knowledge Media," The Journal of the Learning Sciences, vol. 1, pp. 37-68, 1991.

[10] M.Scardamalia and C. Bereiter, "Knowledge building: Theory, pedagogy, and technology, " in The Cambridge Handbook of the Learning Sciences, R. K. Sawyer, Ed. New York: Cambridge University Press. , 2006.

[11] M. Van Amalesvoort, J. Andriessen, and G. Kanselaar, "Representational tools in computer – supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams," Journal of the Learning Sciences, vol.16, pp. 485-521, 2007.

[12] M. Scardamilia and C.Bereiter, “Knowledge building: Theory, pedagogy, and technology, ” in in The Cambridge Handbook of the Learning Sciences, R. K. Sawyer, Ed. New York: Cambridge University Press. , 2006, pp. 97-115.

[13] L. P. Rivard, “A review of writing to learn in science: Implications for practice and research, “ Journal of Research in Science Teaching, vol. 31, pp. 969-983, 1994.

[14] V. Prain, “Learning from Writing in Secondary Science: Some theoretical and practical implications,” Internati/onal Journal of Science Education, vol. 28, 2006.

[15] M. A. McDemott and B. Hand,” A secondary analysis of student perception of non-traditional writing tasks over a ten year period,” Journal of Research in Science Teaching, 2009.

[16] M. Gunel, B. Hand, and V. Prain, “Writing for learning in science: A secondary analysis of six studies,” International Journal of Science and mathematics Education, vol.5, pp. 615-637, 2007.

[17] J. P. Gee, "Language in the science classroom: academic social languages as the heart of school -based literacy," in Crossting Borders in literacy and science instruction: perspectives in theory land practice, E. W. Saul, Ed. Newark, DE: International Reading Association, 2004, pp. 13-32.

[18] J. P. Gee, "Language in the science classroom: academic social languages as the heart of school -based literacy," in Crossting borders in literacy and science instruction: perspectives in theory and practice, E. W. Saul, Ed. Newark, DE: International Reading Association, 2004, pp. 13-32.

[19] J. R. Hayes and L. S. Flower, "Identifying the organization of the writing process," in Cognitive processes in writing, L. W. Gregg and E. R. Steinberg, Eds. Hillsdale, NJ.:Erlbaum, 1980, pp. 3-30.

[20] M. Warschauer and P. Ware, "Automated writing evaluation: defining the classroom research agenda," Language teaching research, vol. 10, pp. 157-180, 2006.

[21] M. D. Shermis and J. Burstein, "Automated essay scoring: A cross-disciplinary perspective," New Jersey: Lawrence Erlbaum Associates, 2003.

[22] P. F. Ericsson and R. H. Haswell, "Machine scoring of student essays: Truth and consequences," Logan, Utah: Utah State University Press, 2006.

[23] R. A. Calvo and R. A. Ellis, "Student conceptions of tutor and automated feedback in professional writing.," Journal of Engineering Education, to appear

[24]J. Villalon, P. Kearney, R. A. Calvo, and P. Reimann, "Glosser: Enhanced Feedback for Student Writing Tasks," in *Advanced* Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on: IEEE Computer Society, 2008, pp. 454-458.

[25] R. Ellis and R. A. Calvo, "Discontinuities in university student experiences of learning through discussions," British Journal of Educational Technology, vol. 37, pp. 55-68, 2006.

[26] R. A. Ellis, C. Taylor, and H. Drury, "University student conceptions of learning through writing.," Australian Journal of Education, pp. 6-28, 2006.

[27]J. B. Biggs, Teaching for quality learning at university: what the student does. Buckingham: Open University Press., 1999.

[28] M. Liu, R. A. Calvo, and V. Rus, "Automatic Question Generation for Literature Review Writing Support," in Intelligent Tutoring Systems. vol. 1, V. Aleven, J. Kay, and J. Mostow, Eds. Pittsburgh, USA: Springer, LNCS 6094, 2010, pp. 45-54.