R.A. Calvo. Affect-Aware Reflective Writing Studios. (2015). In R.A. Calvo, S.K. D'Mello, J. Gratch and A. Kappas (Eds). Chapter 33. Handbook of Affective Computing. Oxford University Press.

Affect-Aware Reflective Writing Studios

Rafael A. Calvo

School of Electrical and Information Engineering

The University of Sydney

Rafael.Calvo@sydney.edu.au

Abstract

Writing involves complex affective and cognitive processes highly influenced by the environments where we write, and these environments are undergoing critical change. Writing in the 21st century is typically performed using digital devices connected to the Internet, enabling writers to interact with content and with other people in completely new ways. This increased interconnectedness opens the door to new opportunities as well as challenges. This paper proposes a new type of tool, Reflective Writing Studios, that can be used to study writing phenomena in a most encompassing way, taking into account the writer's physical and social surroundings and their emotions and mental states in addition to the cognitive processes generally studied. The paper introduces the architecture for the affect-aware multimodal interaction system, its components and their evaluation. Two types of data are used: structured information about the activity and multimodal sensor data from the writer and the environment. Data fusion and user interaction (e.g. visualizations) are developed using affective computing techniques. We discuss how RWS can support novice writers to reflect on the context, process, and outcomes, which are aspects acknowledged as key for developing writing skills. The framework opens avenues for research in terms of multimodal data collection and interpretation.

Keywords: writing, cognitive technologies, affective computing, ubiquitous computing, reflection

1. Introduction

Writing skills are required to varying degrees in all professional disciplines, not only as a method of communication, but also as a catalyst for higher cognitive functions such as analysis and synthesis (Emig, 1977). Teaching writing is notoriously challenging and the development of writing skills in learners at any level has been acknowledged as one of the biggest challenges in education (*The Neglected "R": The Need for a Writing Revolution*, 2003). The challenges are now compounded by the disruptive changes produced by ubiquitous new technologies. The Internet has become an external memory, changing the ways we remember the facts needed for writing (Sparrow, Liu, & Wegner, 2011); social networking changes the way we work, feel and relate to others (as shown by a myriad of studies, some of which are described in this volume) and technologies designed to grab our attention (e.g. mobile phones) are embedded in the environments that surround us (Vertegaal, 2003). Although writing was rarely done in complete isolation, new technologies are now interjected into all its underlying processes: cognitive, affective and behavioural.

New tools and theories that help us understand the effect of these digital environments, and help writers cope with them is gaining urgency. Even before these new developments researchers acknowledged that writing environments had not helped improve writing skills or learning (Pea & Kurland, 1987), rather focusing on improving productivity and presentation. This limitation was first raised soon after word processors became widely available (Pea & Kurland, 1987), and still applies.

These tools would use that about the writer and her/his environment. Current methods used to collect data about writing processes include: methods to help understand behaviour (screen capture, eye tracking, keystroke logging (Bixler & D'Mello, 2013; Wengelin et al., 2009)), qualitative methods, field and ethnographical studies that help understand the cultural and environmental conditions, natural language processing, and think-aloud protocols that help understand, among other things, the subjective experience of writers. So far, analysing the data collected through these methods requires a researcher's time and expertise, and this means that individual writers do not generally see, and benefit, from the outcome of such analysis.

The data can then be used to scaffold the writer's reflection about the process. Writers make a myriad of complex decisions regarding the composing activity (e.g. topic and argument), the reading, the environment and tools, the timing and so forth. We argue that current sensing and affective computing techniques can be used gather empirical evidence for new writing theories and automated feedback that goes beyond the document's quality features. Learning about these many issues requires reflection, particularly reflection-in-action (Schön, 1983), as a transformative element. The literature on reflection is used as a framework for our proposal on how to build systems that help novice writers develop their skills. The HCI community has used sensor data to create tools that support reflection, for example, in 'personal informatics' (Sas & Dix, 2009). We aim to develop these ideas further and contribute to the design of novel writing support tools.

Our research on Reflective Writing Studios takes into account behavioural and emotional expressions. As often found in descriptions of learning interactions, like dialogues between a coach and student described by Schön (1983), the behavioural and emotional expressions of subjects are key to understanding the writing and reflection phenomena. For example Schön (1983) describes how the learner may express ‘surprise’ at something the coach says or mention the perceived approval/disapproval in the coach’s voice. These emotional variables are being increasingly studied in the literature. Computer-based approaches to detecting emotions and behavioural information are progressing rapidly (Calvo & D’Mello, 2010) and specifically in the context of learning interactions (Calvo & D’Mello, 2011).

In Section 2, we briefly discuss how different perspectives come together to inform the idea of Reflective Writing Studios. These include 1) the cognitive process model of Flower and Hayes (1981); 2) the information processing model of Gailbraith (Torrance & Galbraith, 2006), 3) theories that focus on the social aspects or on the materiality of writing (Haas, 1995); 4) the reflection and modelling processes of Schön, and those that study the affective/emotional aspects of writing. In Section 2 we also discuss a series of questions that a writer can reflect on. These include conceptions and motivation, the process and outcomes of the activity, the environment, mental states and ancillary activities that arise during the activity. In Section 3 we describe the system architecture of our Reflective Writing Studios and review four projects that form the key components of the system. The first one, iWrite, focuses on managing the writing activity (deadlines, topic and genre, group members - if collaborative), integrating tools for peer and tutor feedback, and providing disciplinary specific tutorials for the support of writing in different genres. This component covers the formal structure of the activity (i.e. what the tutor had planned). The second component represents an ongoing effort to build systems that collect extensive information about the computer-human interaction taking place while the user is writing. This component covers what actually happens during the activity, the aspects that the instructional design did not consider, for example when the writer engages in extracurricular activities (e.g. reading email). The third component is the engine that fuses together the information sources from the above two. The fourth component provides the user interface with questions that scaffold the reflection and interactive visualizations generated from the data analysed. In Section 4 we discuss the results and forthcoming challenges.

2. Theoretical perspectives

We ground our research program in the extensive literature at the intersection of writing research, psychology and human-computer interaction. The theoretical models used to understand writing processes and the technologies that afford functionalities suggested by these models maybe useful to researchers in other application domains but comprehensive descriptions of these theories are out of the scope of this chapter. We recommend, for example, the Handbook of Writing Research (MacArthur, Graham, & Fitzgerald, 2006) and the Sage Handbook of Writing Development (Beard & Riley, 2009). The most widely used models of writing since the 1980's are probably the cognitive models of writing (Flower & Hayes, 1981) which inform the design of functionalities found in today's word-processing tools such as tools for writing outlines, online thesaurus, and spelling checkers. Debates about the human information processing limits and dual task interference have driven research into attentive computing. Socially situated models of writing were proposed to address the cognitive model's emphasis on the solitary author (Cooper, 1986). These models led some HCI researchers to develop social-awareness tools (Dourish & Bellotti, 1992). Others have focused on the physical aspects of writing e.g. the actual places and situations where writing takes place (Haas, 1995). HCI and ubiquitous computing researchers also use sensors to study the embodiment of cognition and how behavioural (Dourish, 2004) and physiological (Sharma, Pavlovic, & Huang, 1998) information can be used to improve computer interactions.

Grounded in a cross-section of these theoretical frameworks we have worked on supporting aspects of writing, and issues that writers could reflect on, together with specific functionalities that Reflective Writing Studios provide. Some of these elements arise before the writing activity has even begun and are driven by the literature on relational student learning (Marton & Booth, 1997) and motivation:

1. What is the purpose of this activity? (*Conceptions*).
2. How much effort do I need to invest? How likely am I to succeed? (*Metacognition and Motivation*).

Other questions are more closely related to the cognitive model and relate directly to the process and the outcomes of writing:

3. Have I addressed the requirements of the assignment? Covered the topics? Addressed the quality measures by which I will be assessed? (*Outcomes*)
4. How will/did I complete the assignment? Should I Google this term? (*Process*)

A third group of questions refers to the context of the cognitive activities, the environment, and ancillary activities that surround the writing:

5. Should I work at home or in the café? Should I turn down the music? Should I turn off IM? (*Environment*)
6. Am I tired? Bored? Did I enjoy this task? (*Affective states*)

7. Should I take a break and read my email or Facebook? (*Ancillary activities*)

2.1 Conceptions

Relational student learning research (Marton & Booth, 1997) has shown how students' conceptions of a task influence what they do and what they learn. This body of research has provided particularly strong evidence that the way students think about a task, and the few qualitatively distinct strategies they adopt, are closely related to their levels of achievement (Calvo & Ellis, 2010). Studies (c.f. Calvo & Ellis, 2010) have shown that writers have a few (2 - 4) qualitatively different conceptions of writing (e.g. surface vs. deep learning), and that these variations affect their learning outcomes. These views are also similar to those in other studies that explore the conceptions of writing in college students (Hounsell, 1997; Lavelle, 1993). What is of particular importance about this literature is that, at a certain level, the computer interfaces do not have to be particularly intelligent to help writers. This is because since conceptions are so important, writing systems can be augmented to include content emphasizing the purpose (e.g., learning) and the value of both writing and feedback (Calvo & Ellis, 2010). This is of particular importance when writing is not an acknowledged core skill (e.g. in the engineering disciplines) so students might have widely varying conceptions about its purpose.

2.2 Motivation

There is extensive psychological research on the socio-cultural influences on motivation and learning, and comprehensive surveys of the literature (e.g., McInerney & Van Etten, 2004). Some of these models of motivation have been extensively evaluated. One example is Keller's (1987a) ARCS model, which assumes that an individual will be motivated (i.e. engaged) with an activity in proportion to how much *value* (something to gain from it) (s)he perceives there is and in proportion to the perceived likelihood of being successful (the *expectancy* component). The model defines four dimensions of motivation: Attention, Relevance, Confidence and Satisfaction.

Although the literature provides a number of models like ARCS for understanding the major influences on the motivation to learn, the understanding that students have about their own motivation, or the effect of reflecting on this, has not been as frequently studied. Of relevance to the discussion here is that these metamotivational drivers may be more relevant than the motivation itself. As example of how motivational drivers can be used by reflective writers, this chapter considers the first condition of the ARCS model. Sustaining students' attention during an activity is a key concerns for instructional designers. According to the

ARCS model, novice writers would benefit by reflecting on how certain external events, for example, external stimuli (e.g. environmental noise) affects their attention. This can be achieved using the type of sensing devices and software discussed in the next section.

2.3 Outcomes.

Once the writing task is completed, writers need metacognitive skills to assess the quality of the final product. Students should reflect and deepen their understanding of how well their writing fulfils expectations.

The Glosser system, discussed in more detail later, provides several tools for helping writers identify the topics they have covered, for example through concept maps (Villalon & Calvo, 2011), and whether those topics form a coherent argument (O'Rourke & Calvo, 2009b). These visual and text representations are complemented with trigger questions such as: *Could the ideas in the essay be organized in a clearer way*? *Does the composition address the topics requested*?

2.4 Process

Reflection on questions such as 'How will/did I complete the assignment?' can help develop meta-cognitive understanding on the process followed. Log files can be analysed to produce process models (Van der Aalst, Weijters, & Maruster, 2004), an approach that is increasingly popular in educational data mining (Baker & Yacef, 2009; Romero & Ventura, 2007). In the context of teaching writing, this type of feedback is being provided in the WriteProc project (Southavilay, Yacef, & Calvo, 2010) described in more detail later. Most writing activities are assessed on the product (i.e. the student's submission) rather than on the process. It should then not be a surprise that students mostly focus on the document they must submit. Their time-on-task and their reflection would not likely be on the process they followed.

When a writer reflects on the process by which s(he) got to certain point or achieved certain outcomes, the reflection can help identify the important cognitive (e.g. in depth understanding of the topic), affective (e.g. engagement) and behavioral (e.g. longer writing sessions) patterns. Feedback does not generally promote process-oriented reflection either. In the classroom, students and instructors are accustomed to receiving feedback on the assignment that was submitted, but not often on what the student did in order to get to that outcome. For example, if a team working on a proposal for an innovative product does not engage in a brainstorming session and focuses on the first idea that came to a team member's mind, the outcome might not be as innovative and the learning experience not as valuable.

Increased focus on the process can have other benefits. For example, since plagiarism, a common problem in writing assignments, is a means to an end (i.e. to skip the process that requires effort), one could argue that the focus on outcomes is one main factor that drives students to plagiarize. If the focus was more on the process inappropriate behaviours could be reduced.

Most automated feedback tools, and all those that focus on summative assessment only analyze the writing product. Technically providing feedback on the process of writing is only possible when the document's history is stored and mined. Writing processes can be studied to improve the effectiveness of an individual or team. Fortunately, there are modern day writing tools (e.g. Google Docs) that record the revision history in a way that we can use in feedback systems.

2.5 Surroundings: Physical and social

Ede and Lundsford (Ede & Lunsford, 1992) showed that 85% of the documents produced in offices and universities had at least two authors. This line of research has raised the awareness on the impact of social environment on writing activities. Most of the progress has come from the computer supported collaborative work area which provides new forms social awareness and collaboration tools. Interestingly, despite the importance of learning to write as part of a group, students tend to dislike these activities, or consider them as requiring more effort than they are worth (Aditomo, Calvo, & Reimann, 2011).

The systems we, and other affective computing researchers, are developing combine sensor data about the social and physical context in which an activity happens. The HCI community is increasingly interested in building tools for reflecting on data collected through sensors and self-reports (Li, Forlizzi, & Dey, 2010), but as far as we know none has been on writing.

3.6 Emotions and mental states

There has been a surge of research on affect among learning technologies (R. A. Calvo & D'Mello, 2011). Until now, the focus of most research had been on issues such as test anxiety (Pekrun, Goetz, Titz, & Perry, 2002; Schutz & Pekrun, 2007), but new affective computing and sensor technologies allow computer models that can detect affect automatically. They require self (by the subject) or expert reports to be collected, arguably for each subject, and used to train supervised models. These models then 'recognize' the mental state from facial expressions, physiology, posture or other expressive and behavioural channels, some no more invasive than a webcam. Visualizations of such states, for example a view of mental states across the timeline, can provide significant avenues for reflection.

3.7 Ancillary activities: distractions and breaks.

Writing activities often require many hours of work. Writers must stage their work in phases between which rest and other activities occur. These activities will influence the way we approach the writing task and its outcomes. If we do not rest when we are tired we cannot be productive, but if the rest time is used for other cognitive activities this may not have the desired effect. On the other hand if rest becomes procrastination, these interruptions might not be as fruitful either.

3. System Design

Our Reflective Writing Studio is made of four components shown in Figure 1. Each component, and its evaluation(s) is described in the following subsections. Each of the components maps onto a set of requirements:

1) Structured information about the activity (e.g. deadlines and genre) are provided by a writing activity management system (iWrite (Rafael A. Calvo, O'Rourke, Jones, Yacef, & Reimann, 2011) described later). This component also manages structured content that learners can use for modelling (e.g. sample documents) and that the system can use to infer the topics automatically. The information is generally provided by the instructor.

2) The system also needs to collect contextual information - including information not specified by the instructor – about the user and the environment using sensors and ubiquitous computing methods, not included in the original specification by the instructor. In our system this is to be performed by WebEmo (Aghaei Pour & Calvo, 2011), a browser extension that records multimodal information (e.g. from a webcam) together with the actual content written.

3) Machine learning techniques are used to process and integrate the multimodal signals coming from the sensors and the structured information from the activities (in our implementation this is provided by the Siento module).

4) The aim of the RWS is to provide feedback to writers. This is done through a framework that uses the processed data to provide automated or human feedback and scaffold reflection. In our writing environment this would be an extension of the current Glosser project.

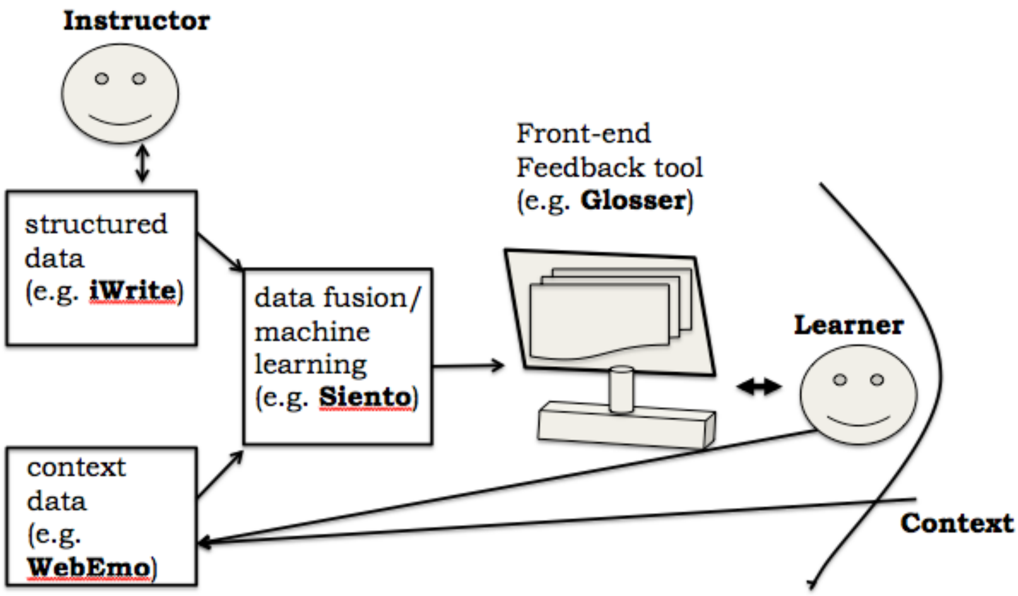

Figure 1: Reflective Writing Studio components

3.1 Writing activity management system

A reflective writing studio requires knowledge about the context of the activity to provide meaningful feedback. Our iWrite (Rafael A. Calvo et al., 2011) writing activity management system, uses genre, deadlines, and the nature of the activity (e.g. individual vs. collaborative) to customize feedback. iWrite is a web-based application that supports individual and collaborative writing. For students, it is the main point of access to the writing activity, collaboration, and instructor provided resources. It also allows researchers and instructors to learn more about the student writing activities, particularly about features of individual and group writing activities that correlate with quality outcomes. It leverages mainstream writing tools (e.g. Google Docs) within which students do the actual writing and facilitates the provision of human (peer and tutor) and automated feedback. iWrite is currently used to support the teaching of academic writing at the Faculty of Engineering and IT at the University of Sydney to around 600 undergraduate and postgraduate students each semester.

The instructional feedback provided aims to advance the students' conceptual understanding, beyond surface features of writing and grammar, to greater awareness of the writing process and textual practices.

The system provides these innovative elements:

- Features to manage writing activities in large cohorts, particularly the management and allocation of groups, peer reviewing and assessment.

- Combines synchronous and asynchronous modes of collaborative writing.
- Uses of computer-based process discovery methods to provide additional information on the team process. The combination of these methods with text mining is particularly novel and will allow feedback about the team processes, based not only on events, but on their semantic significance.

The system uses Google Docs revision management and APIs to record detailed information on the process of writing. In a recent study (Rafael A. Calvo et al., 2011), involving 491 students who completed 642 individual and collaborative writing assignments, 102,538 revisions representing over 51,000 minutes of students writing work were recorded (revisions are saved only when a change is made). The revisions of a document are used to identify different 'activities' (e.g. editing, outlining) that students engage in at different stages of the writing process (Southavilay et al., 2010). These activities are used to produce visualizations that show writers what they have done and aid reflection about which sequence of activities leads to better outcomes.

The revision data also allows for inferences and reflection on time-management issues. Analysis of the data (Rafael A. Calvo et al., 2011) suggested that:

- Students with low grades engage in more revisions (individually and as a group) compared to those with medium grades.
- Students with low and medium grades engaged in fewer writing sessions compared to students with high grades.
- Students with low and medium grades engaged in fewer writing days compared to students with high grades.

Students who obtained high grades were in teams that engaged more frequently in sustained writing sessions. In contrast, shorter burst of document revisions were associated with lower grades. These results seem to agree with those documented by Torrance and Galbraith (2006).

We are planning an integration of iWrite with two Learning Management Systems (Moodle and Blackboard) to collect information about prior experiences of the student. This will be used to provide a user model to generate better feedback.

3.2 Recording platform

Reading, writing and collaborative activities are generally computer-based. The writer can engage in the prewriting activities (e.g. research on the topic to write about), collaborate with peers and tutors, and do the writing itself, all on a browser connected to the Internet. Given the central role of the browser, we used it as a platform called WebEmo (Aghaei Pour & Calvo, 2011) to manage the collection of environmental information while a person reads and writes.

The tool consists of several components shown in Figure 2. The client-side component works as an extension in Google's Chrome browser and does not require any code in the web page to be altered. Through integration with other systems it currently records physiological signals such as ECG (heart), EMG (facial) and respiration patterns. It also records video of the user currently from a webcam and using the openCV library. The actual recording is left to third party software managed by an *acquisition system controller* allowing the framework to be extended to include other modalities. All data collected from each sensor (different physiological modalities and camera) is stored locally and uploaded to the processing server at the end of the session. Video recording can be done with any webcam or with a range camera like Microsoft Kinect. For physiological recording, WebEmo integrates in-house and commercial systems. We generally use the BIOPAC MP150 system with AcqKnowledge software to capture data. Biopac's Matlab API can also be used for real-time data acquisition. The Biopac system can be used to record a variety of physiological signals (ECG, GSR, EMG, EEG etc).

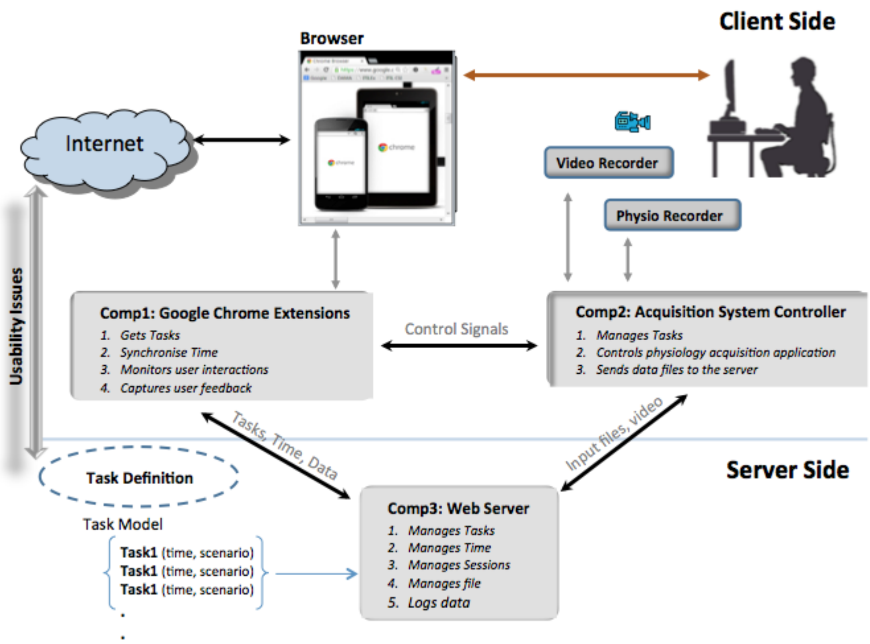

Figure 2: WebEmo architecture for recording users' video, physiology and web interaction data.

When self-reports of the affective impact of a page or activity need to be recorded, the framework can intervene with psychometric questionnaires such as the Self Assessment Mannequin (Lang, Greenwald, Bradley, & Hamm, 1993).

Some studies, those which examine writing in complex realistic settings, would require all of the user's activities to be recorded (perceptions might be affected by what happened immediately before or after the task). For example, a student writing an activity on Google Docs would visit Wikipedia for information on a particular topic, so when testing such a system, the interaction on both websites needs to be recorded. A challenge for the system is that the user can be interacting with different websites running on different browser tabs.

3.3 Data fusion and processing

The systems above can be used to collect vast amounts of multimodal information on what writers do, think, feel, and on the environment that surrounds them. The data is complex, highly dimensional and dynamic, a situation increasingly common in ubiquitous computing scenarios. Researchers of different fields are trying to make sense of this data.

In the area of emotion detection, 'affective computing' techniques (R. A. Calvo & D'Mello, 2010) are being used. These methods examine data recorded from different modalities (sensory features) to train computational models to recognize patterns that correlate with what people report as emotional states. Facial expressions and speech and metalinguistic information were among the earliest types of data to be analysed; currently behaviour, posture, breathing and physiology are all commonly used in affective computing. Many researchers believe that computers that aim to have detection accuracy similar to that of humans should also use multimodal approaches (Sharma et al., 1998).

We have built Siento, a system to record and process multimodal signals and predict affective states. The tool has been used in several affective computing studies combining physiological signals with facial expressions. The system allows for dimensional or categorical models of emotions, self-reported vs. third party reporting, and can record and process multiple types of modalities including video, physiology and text. This type of system can improve the repeatability of experiments. The system is also used for data acquisition, feature extraction, and data analysis consisting of applying machine learning techniques.

3.4 Automated feedback

Glosser is a web-based framework for providing automated feedback on writing (R. A. Calvo & Ellis, 2010; Villalón, Kearney, Calvo, & Reimann, 2008). Any version of a document can be processed to produce a wide range of feedback on collaborative or individual writing activities spanning visual and text modalities. Feedback can be on surface or content features, on the writing product (the final document), or on the process.

For each activity one or more forms of feedback can be selected by the instructor. Each form of feedback, such as the one shown in Figure 3, is made of 1) a set of guiding questions aimed at scaffolding reflection on particular aspects of the writing; 2) a 'gloss', a visual or text-based representation of automatically generated feedback specific to the student's composition; 3) guiding text explaining on how to use the gloss to reflect on the questions. The architecture incorporates features for feedback forms such as argument quality, textual features (such as coherence), automatic generation of questions and feedback on the process. Computer-based text analysis methods are used to provide additional information on text surface level and concept level to writing groups.

A number of feedback tools have been built using the framework. These include two for helping writers reflect on flow: one using interactive text and another a map (R. A. Calvo & Ellis, 2010; Villalón et al., 2008). Three tools support reflection on the topics and concepts: two visual representations using different computational approaches and a textual one. One tool is for reflecting on the social and time related aspects (i.e. the process) (O'Rourke & Calvo, 2009a) and one generates questions automatically (Liu, Calvo, & Rus, 2010).

The automatic question generation techniques (Liu et al., 2010)(Liu, Calvo, Aditomo, & Pizzato, 2012) are used to support students to reflect on different aspects of their literature reviews. The system scaffolds students reflection on their academic writing with content-related trigger questions automatically generated (using NLP techniques) from citations. A taxonomy of different types of citations relevant to literature review papers (Liu et al., 2010) is used (with categories for Opinion, Result, Aim of Study, System, Method, and Application). For example, if a student (citer) cites the following opinion in his review: "Cannon (1927) challenged this view mentioning that physiological changes were not sufficient to discriminate emotions", the system can generate the following trigger questions that help the writer reflect on the relevance to the project at hand: *Why did Cannon challenge this view mentioning that physiological changes were not sufficient to discriminate emotions? How is this relevant to your project?*

The system was evaluated in a study with 57 participants (33 PhD student writers and 24 supervisors), where each PhD student submitted a research proposal to iWrite, our writing management system. Each proposal was read by a peer and the PhD supervisor, both of whom provided feedback in the form of reflection questions. The peer and supervisor questions were compared to the automatic questions and also to a set of generic questions. A total of 615 questions were generated based on the 33 literature review papers. Each student was then asked to rate the quality of questions received on measures of ‘acceptability’ (whether it is grammatically correct, not vague, and makes sense according to the context) and ‘usefulness’ (whether it is helpful to trigger reflection). These ratings were used to evaluate the system’s performance and to analyze human expert generated questions. We compared automatically generated questions with human-generated and generic questions on these dimensions using a Bystander Turing test and the top questions (selected using an automated ranking system) were reported to be as useful as those generated by peers and supervisors (after removing questions with surface errors).

The computational techniques (e.g. Latent Semantic Analysis) used in Glosser are similar to those in automatic essay assessment systems. Criterion (by ETS Technologies), MyAccess (by Vantage Learning) and WriteToLearn by Pearson Knowledge Technologies are all commercial products increasingly used in classrooms. These programs sometimes provide an editing tool with grammar, spelling, and low-level mechanical feedback. Some of them can also be integrated with university systems using Application Programming Interfaces (APIs). Some provide resources such as thesaurus and graphic features, many of which would be available in tools such as MS Word. Glosser is distinct in several ways including that it is the only one that provides support for collaborative writing activities and the only one that can analyse (and provide feedback on) the writing processes.

With regards to process, recent work such as our Writeproc system (Southavilay et al., 2010), a component of Glosser, shows that the writing processes can be analysed by detecting writing ‘activities’. WriteProc uses a taxonomy of writing activities proposed by Lowry et al. (Lowry, Curtis, & Lowry, 2004). We plan to study how such information can help students reflect on their own processes. If the results of these observations can be generalized to other learning situations, students may benefit from the realization of how important it is to work on an assignment for dedicated stretches of time, and this is a key outcome of learning time management skills.

We have also evaluated how students’ conceptions of automated feedback such as Glosser affect the way they use it, and what the quality of the learning outcomes, i.e. the grades obtained is (R. A. Calvo & Ellis, 2010).

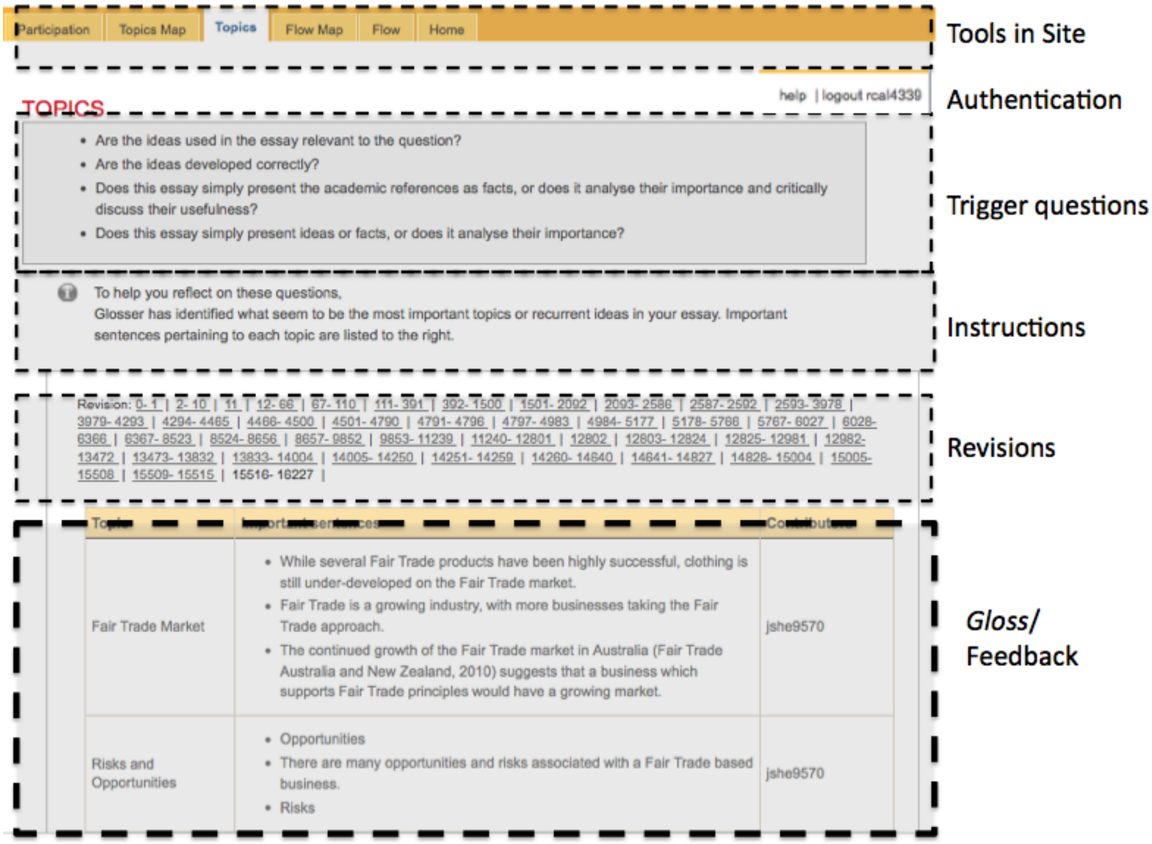

Figure 3: Glosser feedback

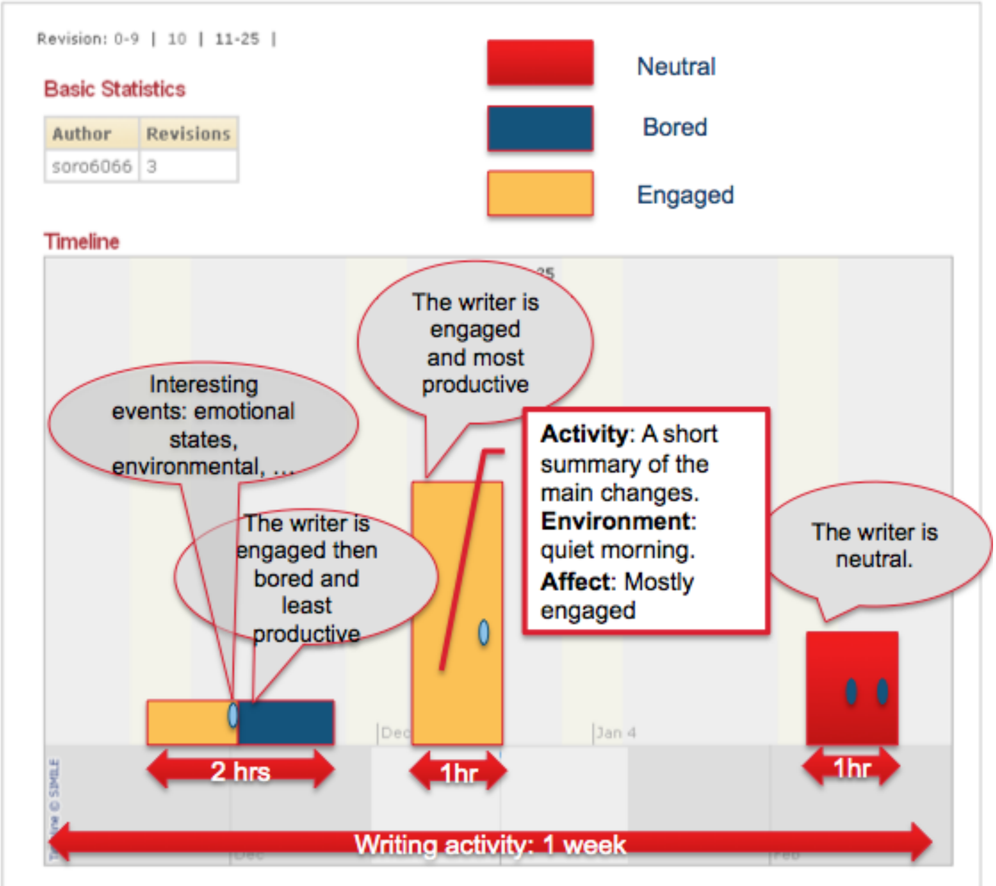

Figure 4: Reflective Timeline

4. Discussion and Conclusions

Each of the four components of our Reflective Writing Studios has been evaluated in the studies described earlier. The integration of these software/hardware components is a significant engineering challenge, so an evaluation of its impact on learning is still some time away. As a way of discussing some of these challenges, I provide an account of how I personally used the system. In a recent session, I wrote a blog post during a 1hr session. The Kinect camera recorded my facial expressions, voice and environmental situation, the writing tool (used by iWrite) was Google Docs that keep a history of the document (a very simplified version of keystroke logging), WebEmo recorded my interactions with the browser. The purpose of the blog post was to summarize a research project of which I was learning about. During the session, a phone call interrupted my work, and I had to stop when someone came in the room. I was following a think-aloud protocol so my speech could be automatically transcribed using a speech recognition system and key events in the writing process could be identified. Speech recognition systems are not 100% accurate but produce a stream of words that can be used by text mining algorithms. The same happens with facial expression recognition. About 20 behavioural events were identified in the video by an annotator. These included affective states, such as surprise and engagement that we have detected automatically, typing and 'habits' (e.g. playing with my ear). My posture was used to detect when a phone call and a visitor interrupted my work. My research (e.g. reading webpages about the project) and writing strategies (e.g. copy and paste of certain details) were all recorded. All this information can be used as evidence in writing research studies and to generate visualizations (integrated into Glosser) that support reflection.

Figure 4 shows a timeline of significant events, such as the ones above, to help a writer reflect on positive and negative aspects across multiple sessions. The height of the bars represents a measure of productivity such as the number of text revisions, addition or deletions, over each time period, and the width is the amount of time averaged over 30min or 1hour blocks. The colour of the bars represents the most common automatically detected mental state, and the square dialog box (that appears on a rollover) provides a description of the block. This description summarizes the structured information coming from iWrite with process data. Small markers can appear on the timeline (e.g., in the first block) highlighting specific meaningful events (e.g., a phone call that may have produced distraction). The curved bubbles are explanatory material to support new users of the interface.

There are two sets of challenges for reflective writing studios. First are the technical issues of building a reliable, accurate, and scalable platform. The second is to produce visualizations and feedback that summarize the massive amount of data being collected in a way that is meaningful to the researcher and/or reflective user.

Automated feedback systems have focused on the writing product rather than the process, for example those based on automated essays scoring. It is not yet clear that students benefit from seeing how their writing products measure against certain scales, but the results are still useful to inform educational design decisions. We speculate that the relationship of such systems to student learning would be stronger if systems focused on process factors such as decisions on topic and argument, supporting reading activities, the selection of environment and tools, the timing, and so forth. The pervasiveness of computing devices, the easy access to supporting material and social interaction, makes many factors that have not previously been included in the study of writing phenomena potentially influential to the way we think, feel and write.

As a consequence, this paper argues that the writer's affect and his or her environment must also be taken into account when designing systems that support writing. With rare exceptions, this has not been the case. One of the reasons researchers have not taken this inclusive approach is that such a complex system is difficult to study using traditional methods.

It is not possible to control for all possible variables occurring in the mind, body and world around the writer.

The Reflective Writing Studios described here are an example of an affective computing application aiming to support novice writers through data-driven functionalities that scaffold their reflection. The functionalities described can help novice writers reflect on the situations when they enjoyed and were productive writing, answer questions such as *why* they engaged in the activity, *how* they went about it, *what* were the outcomes are (so they can be improved on), and *who* did what (when working in a team). This paper first reviewed the literature coming from different pedagogical traditions, including cognitive, social, situated and embodied cognition perspectives and argued for approaches that take into account a wider set of variables. It then argued that new enabling technologies such as affective and ubiquitous computing allow researchers to develop new type of interfaces that can support new forms of reflective feedback.

We have described the different components of the system and their evaluations: the writing activity management system (currently used by over 1,000 students at the university of Sydney) manages activities were writers use Google Docs, and structured data about the activity is stored. Since most of the writing and the ancillary reading activities happen on the browser, the recording platform was built as an extension of Google Chrome, and integrates the webcam (with video and sound), screen recording and other sensing devices. The data collected by the extension and its devices is then merged and mined. We have summarized the evaluations of these components. The final component is the user interface used to provide feedback to the student. We are building this as an extension to Glosser, an open source feedback framework, which we have already described and evaluated elsewhere. A timeline visualization was proposed, but not yet evaluated, to represent the data provided to writers.

Due to the significant amount of information that Reflective Writing Studios collects, such as data about the writing process, it is important that writers do not get distracted or feel overwhelmed by it. It is also important that they do not see it as surveillance of their activities. They must feel in control of their data. An approach we have used for tackling this challenge is to make anonymous all process data shown to instructors. With this setup, students can get feedback at any stage of their writing process but know that instructors only access the final product (as occurs traditionally).

Acknowledgements

Assignment Manager was funded by a University of Sydney TIES grant. Glosser was funded by Australia Research Council DP0665064, AQG and WriteProc were funded by DP0986873, and a Google Research Award.

References

Aditomo, A., Calvo, R. A., & Reimann, P. (2011). Collaborative writing: too much of a good thing? Exploring engineering students' perceptions using the Repertory Grid. *Computer Supported Collaborative Learning - CSCL*. The University of Hong Kong, Hong Kong, China.

Aghaei Pour, P., & Calvo, R. A. (2011). Towards automatic measurements of web usability using affective computing techniques. *Affective Computing and Intelligent Interaction* (p. to appear). Memphis, USA: Springer LNCS.

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–17.

Beard, R., & Riley, J. (2009). *The SAGE Handbook of Writing Development*. London, UK: Sage Publications.

Bixler, R., & D'Mello, S. (2013). Detecting boredom and engagement during writing with keystroke analysis, task appraisals, and stable traits. *Proceedings of the 2013 international conference on Intelligent user interfaces* (pp. 225–234). ACM.

Calvo, R. A., & D'Mello, S. K. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. doi:10.1109/T-AFFC.2010.1

Calvo, R. A., & D'Mello, S. K. (2011). *New Perspectives on Affect and Learning Technologies* (p. 284). New York: Springer.

Calvo, R. A., & Ellis, R. A. (2010). Student conceptions of tutor and automated feedback in professional writing. *Journal of Engineering Education*, *99*(4), 427–438.

Calvo, Rafael A., O'Rourke, S. T., Jones, J., Yacef, K., & Reimann, P. (2011). Collaborative Writing Support Tools on the Cloud. *IEEE Transactions on Learning Technologies*, *4*(1), 88–97. doi:10.1109/TLT.2010.43

Cooper, M. (1986). The Ecology of Writing. *College English*, *48*, 364–375.

Dourish, P. (2004). *Where the action is: the foundations of embodied interaction*. The MIT Press.

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of the 1992 ACM conference on Computer-supported cooperative work* (pp. 107–114). ACM.

Ede, L. S., & Lunsford, A. A. (1992). *Singular texts/plural authors: Perspectives on collaborative writing*. Southern Illinois Univ Pr.

Emig, J. (1977). Writing as a Mode of Learning. *College Composition and Communication*, *28*, 122–128.

Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, *32*(4), 365–387.

Haas, C. (1995). *Writing Technology: Studies on the Materiality of Literacy*. Mahwah NJ: Lawrence Erlbaum.

Hounsell, D. (1997). Contrasting conceptions of essay-writing. *The experience of learning*, 106–125.

Keller, J. M. (1987a). The systematic process of motivational design. *Performance+ Instruction*, *26*(9-10), 1–8.

Keller, J. M. (1987b). Strategies for stimulating the motivation to learn. *Performance+ Instruction*, *26*(8), 1–7.

Lang, P. J., Greenwald, M., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Evaluative, facial, visceral, and behavioral responses. *Psychophysiology*, *30*(3), 261–274.

Lavelle, E. (1993). Development and validation of an inventory to assess processes in college composition. *British Journal of Educational Psychology*.

Li, I., Forlizzi, J., & Dey, A. (2010). Know thyself: monitoring and reflecting on facets of one's life. *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* (pp. 4489–4492). ACM.

Liu, M., Calvo, R. A., & Rus, V. (2010). Automatic Question Generation for Literature Review Writing Support. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent Tutoring Systems* (Vol. 6094, pp. 45–54). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-13388-6

Liu, M., Calvo, R., Aditomo, A., & Pizzato, L. A. (2012). Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support. *IEEE Transactions on …*, *5*(3), 251–263.

Lowry, P. B., Curtis, A., & Lowry, M. R. (2004). Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice. *Journal of Business Communication*, *41*(1), 66.

MacArthur, C. A., Graham, S., & Fitzgerald, J. (Eds.). (2006). *Handbook of writing research*. New York: The Guilford Press.

Marton, F., & Booth, S. (1997). *Learning and Awareness*. Lawrence Erlbaum Assoc, Publishers.

McInerney, D. M., & Van Etten, S. (Eds.). (2004). *Big Theories Revisited* (p. 367). Grenwich, Connecticut, USA: Information Age Pub Inc.

O'Rourke, S., & Calvo, R. (2009a). Semantic visualisations for academic writing support. *14th Conference on Artificial Intelligence in Education* (pp. 173–180). Brighton, UK: IOS Press.

O'Rourke, S. T., & Calvo, R. A. (2009b). Analysing Semantic Flow in Academic Writing. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Artificial Intelligence in Education - AIED* (pp. 173–180). Brighton, UK. doi:10.3233/978-1-60750-028-5-173

Pea, R. D., & Kurland, D. M. (1987). Cognitive technologies for writing. *Review of research in education*, *14*, 277–326.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Emotions in Education: A Special Issue of Educational Psychologist*, *37*(2), 91–105.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *33*, 135–146.

Sas, C., & Dix, A. (2009). Designing for reflection on experience. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4741–4744). Boston: ACM.

Schön, D. (1983). The reflective practitioner. *Temple Smith, London*.

Schutz, P. A., & Pekrun, R. (2007). *Emotion in education*. Academic Press.

Sharma, R., Pavlovic, V. I., & Huang, T. S. (1998). Toward multimodal human-computer interface. *Proceedings of the IEEE*, *86*(5), 853–869.

Southavilay, V., Yacef, K., & Calvo, R. A. (2010). Process Mining to Support Students' Collaborative Writing. In A. Merceron, P. Pavlik, & R. Baker (Eds.), *Third International Conference on Educational Data Mining (EDM2010)*. Pittsburgh, USA.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *science*, *333*(6043), 776.

The Neglected "R": The Need for a Writing Revolution. (2003). College Entrance Examination Board - National Commission on Writing in American Schools and Colleges. New York.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerlad (Eds.), *Handbook of writing research* (pp. 67–82). New York: Guilford Press.

Van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *Knowledge and Data Engineering, IEEE Transactions on*, *16*(9), 1128–1142.

Vertegaal, R. (2003). Attentive User Interfaces. *Communications of the ACM*, *46*(3), 30–33.

Villalon, J., & Calvo, R. A. (2011). Concept maps as cognitive visualizations of writing assignments. *Journal of Educational Technology and Society*, *14*(3), 16–27.

Villalón, J., Kearney, P., Calvo, R. A., & Reimann, P. (2008). Glosser: Enhanced Feedback for Student Writing Tasks. *2008 Eighth IEEE International Conference on Advanced Learning Technologies* (pp. 454–458). IEEE. doi:10.1109/ICALT.2008.78

Wengelin, A., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods*, *41*(2), 337–351.