An Investigation into Sentiment Analysis and Morale Visualisation in a Digital Backchannel System

by

Peerumporn Jiranantanagorn

A thesis presented to
Flinders University
in total fulfilment of the requirements for the degree of
Doctor of Philosophy

Abstract

This thesis investigates the impact of students' sentiments and emotions on learning and teaching in a large lecture environment by using a digital backchannel system. In a large traditional lecture, interactions between a lecturer and students are restricted by many factors, such as seating arrangement and time constraint. Limited teacher-students interactions make it difficult for the lecturer to understand students' feedback that may help improve her/his teaching. This problem is more prominent if the lecture is delivered online because the lecturer cannot physically interact with the students.

One of the solutions is to deploy a backchannel system, a virtual space in which students interact with the lecturer and other students by asking questions, sharing their thoughts and engage in collaborative activities without interrupting the current discourse. However, the current backchannel systems have not paid much attention to aggregate and present students' feedback to the lecturer in a meaningful way that is easy to digest in a short time.

The proposed solution in this thesis, *ClasSense*, analyses emotions and sentiments in students' messages in real time and presents results in a morale-graph-based user interface, which is composed of a trend of students' emotions and sentiment associated with morale scores and related posts, to the lecturer. So, she/he can know what students are thinking of and respond to students' feedback accordingly. Also, the *ClasSense* system uses a microblogging user interface that allows students to communicate with their lecturers and other students by using short messages and emoticons to express their opinions and emotions during lecture.

Evaluation of the *ClasSense* system shows that lecturers accept and prefer the morale-graph-based user interface to conventional backchannel user interface, which displays only posts in

chronological order. The lecturers like to see the morale graph to get a sentiment trend during the lecture then browse through related posts at a particular minute to investigate in more details. Students also express positively and agree that the *ClasSense* system make their feedback an important part of the class and increase their interactions with the lecturers.

The contribution of this research is in the design of *ClasSense* user interface that is integrated with a customised sentiment analysis algorithm to provide a sentiment and emotion analysis in the context of teaching and learning in university. A further direction for research is to determine how to improve the sentiment analysis module and user interface to better support users with different requirements and in different contexts.

Certification

This thesis contains no material which has been accepted for the award of any other degree or diploma, except where due reference is made in the text of the thesis. To the best of my knowledge, this thesis contains no material previously published or written by another person except where due reference is made in the text of the thesis. As requested under Clause 14 of Appendix D of the Flinders University Research Higher Degree Student Information Manual I hereby agree to waive the conditions referred to in Clause 13(b) and (c), and thus

- Flinders University may lend this thesis to other institutions or individuals for the purpose of scholarly research;
- Flinders University may reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Signed Dated

Peerumporn Jiranantanagorn

We believe that this thesis is properly presented, conforms to the specifications for the thesis and is of sufficient standard to be, prima facia, worthy of examination.

Signed Dated	Signed Dated
Dr Haifeng Shen	Dr Carl Mooney
(Principle Supervisor)	(Co-Supervisor)

Acknowledgements

I greatly and sincerely appreciate Dr Haifeng Shen for his guidances, understanding and support throughout my PhD. His patience, enthusiasm and hard work are inspirational. Also, I would like to thank Dr Robert Goodwin, my first five-year supervisor, for his valuable discussions, advices and assistance, Dr Carl Mooney, my co-supervisor, for his advices and encouragement in the early years of my study, and Dr Kung-Keat Teoh, for his suggestions and assistance in the later years of my study.

For Keith and Pat Kimber, your kindness has given me and my wife a memorable time in Adelaide. I'm really glad to live with you and appreciated for all your help.

Finally, I would like to thank my family for their supports and most importantly, my beloved wife. Her great support, encouragement and devotion have made all these years the best time of my life.

Publications

The following publications are derived from this research.

1. Jiranantanagorn, P., Goodwin, R. and Mooney, C. (2012). Mobile learning in Thai public university: Opportunities and barriers. *in* Proceedings of the 7th International Conference on Computer Science & Education. URL: https://doi.org/10.1109/ICCSE.2012.6295381

2. Jiranantanagorn, P., Goodwin, R. and Mooney, C. (2013). A Proposed Mobile Learning System for Thai Public Universities. *in* Proceedings of the 8th International Conference on Information Technology and Applications.

3. Jiranantanagorn, P., Shen, H., Goodwin, R. and Teoh, KK. (2015). ClasSense: A Mobile Digital Backchannel System for Monitoring Class Morale. *in* International Journal of Learning and Teaching, Vol. 1, pp. 161-167. URL:https://doi.org/10.18178/ijlt.1.2.161-167

4. Jiranantanagorn, P., Bhardwaj, P., Li, R., Shen, H., Goodwin, R. and Teoh, KK. (2015). Designing a Mobile Digital Backchannel System for Monitoring Sentiments and Emotions in Large Lectures. *in* Proceedings of the 24th Australasian Software Engineering Conference, Vol. 2, pp. 141-144. URL:https://doi.org/10.1145/2811681.2824994

5. Jiranantanagorn, P. and Shen, H. (2016). Sentiment analysis and visualisation in a backchannel system. *in* Proceedings of the 28th Australian Conference on Computer-Human Interaction, pp. 353-357. URL:https://doi.org/10.1145/3010915.3010992

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Problem and importance of this study

The most common methods for a teacher and students to interact in the classroom are verbal questioning and answering. The teacher asks students questions to keep them actively involved in lessons. These questions also give students opportunities to express their ideas and evaluate their learning, and allow the teacher to revise her/his instructional method (Morgan and Saxton, 1991). Gagne et al. (1993) also stated that students' responses are important in learning because they provide a means for the teacher to provide informative feedback to students to help them know what to do next. Student feedback can help improve teaching and develop effective learning (Sadler, 2010). In addition, students' emotions and opinions in their feedback during a lecture have a significant impact on both teaching and learning processes (Russell, 2003). If a lecturer can detect and manage information about their students' emotions and sentiments in a lecture environment, it is possible for her/him to know and fulfil students' potential needs better (Ortigosa et al., 2014). Therefore, recognising and responding to students' emotions are crucial to effective learning and teaching in a classroom.

However, in large classrooms there are many difficulties of teacher-students interaction, such as the time for a teacher to be in contact with students is very limited (VanDeGrift et al., 2002; Anderson et al., 2003), and students' opportunities to interact with their teacher are often determined by their seating arrangement (Roth et al., 1999) or academic achievements (Shachar and Sharan, 1994). In short, the large lecture format impedes students from interacting with the lecturer and does not promote the active learning as recommended in undergraduate education (Chickering and Gamson, 1987).

Recently, the proliferation of smartphones among students makes them an outstanding choice of technology to engage students (Wagner, 2005). Smartphones are no longer just the voice communication devices; they are now universal and carried around by most students at most of their time. They have an impact in almost every aspect of their lives, including their university education. Newer generations of students, who have lived with such technologies for the greater majority of their lives, are passionate and dedicated users of such smart mobile devices. In particular, these students enjoy the connectivity and social interaction which occur from the use of these devices and prefer group-based activities (Cobcroft et al., 2006).

Smartphones can be used to support both independent and collaborative learning (Savill-Smith and Kent, 2003). Research showed that using mobile devices, such as smartphones and tablets, in collaborative learning environments might seem to stimulate more knowledge generation and more learning tasks by impelling more motivation than other learning environments (Ryu and Parsons, 2012).

In the last decade, smartphones have been used to develop backchannel systems because of their ability to run applications and their widespread adoption in classrooms (White and Turner, 2011). A backchannel system is a supplementary virtual space in which students inform the lecturer and other students by sharing their thoughts and engage in collaborative activities without interrupting the current discourse (Pohl et al., 2011). It has emerged to make a large classroom more manageable and engaging.

Some of the backchannel systems that are available on mobile devices are Hotseat (Aagard et al., 2010), Backstage (Pohl et al., 2011), ActiveClass (Ratto et al., 2003) and ClasCommons (Du et al., 2012). More recent studies of backchannel systems allow students to use microblogging style to post questions and opinions through smartphones (Aagard et al., 2010; Pohl et al., 2011). Several studies have explored the benefits of using smartphones and PDAs for students to ask questions and give feedback in the lecture session. Mobile Lecture Interaction project (Järvelä et al., 2007), ActiveClass project (Griswold et al., 2004) and the Hotseat (Aagard et al., 2010) and Backstage systems (Pohl et al., 2011) have attempted to promote interaction in the class by allowing students to use the smartphones and PDAs to anonymously submit questions, respond to poll, discuss with friends, vote on questions and give feedback to the lecturer during the lecture session. Also, students can see questions, comments and votes from their peers.

Even though many backchannel systems have been developed, common to these systems is the focus on designing the user interface to make it easy for students to input feedback and read others' posts. However, not much attention has been paid to help the lecturers easily grasp the aggregated feedback from the crowd and respond to the most important concerns students share in common. For example, posts in Hotseat can only be sorted, such as according to most recent, most popular, or most discussed, and the lecturers using Hotseat and Backstage are unable to quickly gauge the mood of the students in the class and adapt their teaching strategies spontaneously to respond to students feedback. Instead, they can only read the responses on these systems post-class and react in future classes. Also, lecturers using ActiveClass and ClasCommons found difficulties in integrating the system into their lectures. In short, current mobile backchannel systems are not capable of providing lecturers immediate and meaningful responses. Likewise, Fies and Marshall (2006) reported that lecturers found a burden from tracing the individual student's response to answer students' questions because the scattered and sparse nature of posts.

The issue is more prominent if a lecture is delivered online as the lecturer cannot physically gauge the mood of students (George et al., 2000) and has to rely on a backchannel system to receive and respond to students' feedback. In recent years, delivering a lecture online has been increasingly popular. A report based on responses from more than 2,500 colleges and universities in the US stated that there was a higher demand for online courses than face-to-face courses, and there was a compound annual growth rate of 19% of students taking at least one online course during fall 2002 to fall 2008 (Allen and Seaman, 2010). They also revealed that the number of online learners is almost more than the US total higher education student population. One of the latest trends in online learning is a Massive Open Online Course (MOOC) (Dabbagh et al., 2016; Martin, 2012). These massive online learning sites such as MIT Open Course Ware (OCW) (Cormier and Siemens, 2010; Abelson, 2008), edX and Coursera have become an important medium for self-education. More than 1,000 online courses have been released from three major MOOCs platforms (edX, Coursera and Udacity) and the total number of registrants has reached 10 millions (Pappano, 2012). Many educators believe that MOOCs will change the higher education landscape forever (Shi et al., 2015).

However, one of the challenges is that it can be difficult for the teacher to understand learners' feedback that might help them improve their courses. The current asynchronous online course platforms have added monitoring and analysis tools such as edX's Metrics Tab and Khan Academys Coach monitoring system (Stephens-Martinez et al., 2014). However, these tools only provide a limited set of visualizations of basic quantitative information such as the clickstream that consists of various "click" actions (e.g. "play", "pause", "seek") learners used while watching course videos (Shi et al., 2015).

While interactions in asynchronous online courses, such as e-learning and MOOCs, are primarily through lecture videos and discussion forums, synchronous online courses usually provide real-time interactions through virtual classroom technologies. For example, a whiteboard tool for a teacher to write, draw and highlight on her/his presentation, a survey tool to create polling questions for reviewing of lecture content, a text/audio chat function for students to communicate with their peers and the teacher and a desktop/webcam sharing for the teacher to better support students (Hrastinski, 2008; Skylar, 2009). Even though these features are provided, it is difficult for the teacher to be aware and active regarding interaction and communication with a large group of students McBrien et al. (2009).

As a result, in this research, we propose a new backchannel system *ClasSense* to address the common issues found in the existing lecturer support tools when used in a large lecture context including (1) helping a lecturer to process and respond to the large amount of students' posts effectively, (2) getting an overall picture of students' learning as well as the emotions and sentiments of students, and (3) tracing back the summary of incidents that happened in the lecture in order to improve her/his teaching and inspecting the students' understanding level from their responses.

ClasSense allows students to express their emotions and analyses the sentiments of their posts in real time so that the lecturer can monitor the morale of the student population and respond to the most important concerns students have in common. This is the first work that incorporates both emotion and sentiment analysis into a backchannel system for the purpose of giving the lecturer more insights into students' feedback through their emotions and sentiments and studying the impact of students' morale on learning and teaching in a large lecture.

1.2 Contribution

The aim of this research is to study the impact of students' morale on learning and teaching in a large lecture environment by using a backchannel system that can analyse and report students' sentiment and emotion during lecture.

Our contribution is the *ClasSense* framework including (1) the *ClasSense* lecturer application that helps a lecturer promptly interact with her/his students in an online lecture class by displaying the overall students morale and the real-time top ranked posts so that she/he can choose when and how to respond to these posts and/or adjust her/his teaching accordingly, (2) customisation of SentiStrength, a lexicon-based sentiment analysis software, in the *ClasSense* back-end system to analyse sentiments and emotions in our context, and (3) the *ClasSense* student application designed using a microblogging user interface with explicitly displayed textual emoticons for students to choose to embed into their posts.

Chapter 2

Literature review

In December 2010, creating contents for a wide range of mobile devices was a challenging task because there were so many mobile platforms, such as WAP, Blackberry, Android, iOS and Nokia, and HTML5 technologies were not fully supported in all mobile devices yet. As a result, I thought that it would be a good idea to develop a mobile content authoring tool that helps lecturers, especially in developing countries, to automatically convert an existing e-learning content or create new mobile learning contents for various mobile devices because universities in developing countries have limited resources (technical staff and budget) to help develop mobile learning contents for them and these countries have a very high mobile phone penetration rate. So, I started studying about mobile learning and tools which constitutes the first part of this literature in Section 2.1.

However, in 2013, I started to realise the advancement and affordability of smartphones, and technological challenges of cross-platform mobile application development were minimised with the adoption of HTML5 in newer smartphones. With HTML5 technologies, one with basic knowledge of HTML, JavaScript, and CSS can create apps, games or other contents and deploy them to most mobile platforms by using hybrid mobile app frameworks such as Apache Cordova and Appcelerator Titanium. As authoring contents for mobile devices is no longer an issue with the new HTML5 technologies (Anthes, 2012), my attention shifted to the delivery of online contents to students who access them through their mobile devices. This become the second part of this literature in Section 2.2.

2.1 Mobile learning systems and content authoring tools

A number of definitions of mobile learning have been given in various studies, for example, the use of mobile and handheld devices such as mobile phones, Smartphones, and PDAs in the learning process (Ivanov and Momchedjikov, 2009), the delivery of electronic learning materials that work on a wide range of mobile devices (Quinn, 2000; Ally, 2005; Traxler, 2005) and e-learning that happens on any small and portable device such as PDAs, and cell phones, that we can carry and use for accessing content and interacting with other people in everyday life (Trifonova and Ronchetti, 2003). Research has shown that using mobile devices for educational purposes is a cost efficient method of providing education (Motlik, 2008). Mobile learning could play an important role in supporting lifelong learning (Holzinger et al., 2005). Also, mobile learning could improve student's e-learning experience, engagement and interest by providing more channels to access course content and learning activities anytime and anywhere (Chmiliar, 2010).

Due to a diversity of mobile platforms and their inherited limitations such as screen size, resolution and display colour, creating mobile learning content available on different mobile devices is a problem (Niazi and Mahmoud, 2008; Whattananarong, 2005). As a result, mobile learning content authoring tools were developed as one of the most important components in mobile learning systems to help academics who do not have much technical knowledge to create and distribute mobile content. These tools can be broadly divided into textual and multimedia content authoring tools. Some of the previous studies of mobile learning content frameworks and authoring tools are presented below.

2.1.1 Mobile learning framework and architecture

Sharma and Kitchens (2004) proposed a Flexible Services Architecture using web service technologies to allow students to access university's provided services such as Blackboard, library service, online tests and language translation through various Internet connected devices including mobile devices (Tablet PC, PDA, Wireless Phone and Web TV) and personal computers. However, in order to study offline, students had to download course materials to their mobile devices, which was costly at that time.

Issack et al. (2006) proposed a Mobile-E-Learning adaptive architecture. Users can access information such as courses, activities, questions bank and student profiles on the server using web browsers on their desktop computers or mobile devices (smartphone or PDA). The lecturers can create multiple-choice questions and track students' progress from the provided web tools. The students had options to take a test either online with their desktop computers or mobile devices, or offline by downloading the test application to their mobile devices, doing a test and sending their performance data back to the server. The strength of this system was the supplementation of mobile learning without changing the existing e-learning system. However, the server needed to be configured to support quiz downloading and performance synchronisation from students.

Motiwalla (2007) proposed a mobile learning framework, which combined mobile connectivity concepts such as push and pull mechanism with e-learning concepts. This framework utilized a Short Message Service (SMS), Wireless Access Protocols (WAP) and Wireless Markup Language (WML) to develop a prototype application. The mobile learning environment including customized RSS news alerts, discussion board and chat room was accessible from WAP-supported mobile devices and personal computers. The advantages of this architecture were that users can interact with the system and other users from both mobile devices and PCs, and any changes they made or messages they posted will be saved into the same database.

Traditional mobile learning applications had limitations in terms of high cost of devices and network, limited bandwidth and educational resources (Gao and Zhai, 2010). Cloud-based mobile learning applications were introduced to overcome the limited capabilities of mobile devices. For example, utilizing a cloud with large storage capacity and powerful processing ability to support the use of multimedia learning contents in mobile learning (Saranya and

Vijayalakshmi, 2011). The cloud-based mobile learning applications provided learners with much richer services in terms of information size, faster processing speed, and longer battery life. Moreover, the cloud offers a great opportunity to construct a mobile educational resource library (Chen et al., 2011; Bin, 2011)

Zhao et al. (2010) presented the benefits of combining mobile learning and cloud computing to enhance the communication quality between students and teachers. In this case, smartphone software based on the open source JavaME UI framework and Jabber for clients was used. Through a web site built on Google Apps Engine, students communicated with their teachers at any time. Also, the teachers could obtain the information about students' knowledge level of the course and could answer students' questions in a timely manner.

In addition, a contextual mobile learning system based on Mobile Interaction in Augmented Reality Environment platform (Yin et al., 2009) showed that a cloud based mobile learning system can help learners access learning resources remotely. Ferzli and Khalife (2011) developed a cloud-based education tool that was used to create a course about image/video processing. Through mobile phones, learners can understand and compare different image/video processing algorithms used in mobile applications (e.g., de-blurring, de-noising, face detection, and image enhancement).

2.1.2 Textual mobile content authoring tools

A Context-Sensitive Learning Content Management System has been proposed by (Chu et al., 2004) to support mobile learning by transforming the learning content on a desktop computer to formats suitable for different mobile devices. The system includes four modules: an Authoring Tool, a Content Management System, a Content Storage system and Context-Sensitive Middleware. Lecturers are able to upload documents in Portable Document Format (PDF) or as Microsoft Word Documents (DOC) and Microsoft PowerPoint (PPT) slide-shows, following which the system transforms and customizes the documents into the HTML or Wireless Markup Language (WML) file format for display on mobile devices.

A Mobile Author was developed by (Virvou and Alepis, 2005) for instructors to create and manage Intelligent Tutoring Systems (ITSs), databases of student profiles, the domain to be taught, tests and homework using a computer or a mobile device. The main purpose of ITSs is to provide individualised guidance in any subjects to students. After the ITSs were created and distributed to students, students can use their PCs or mobile phones to access theory and tests in ITSs from anywhere.

Mobile learning tools were developed by (Ivanov and Momchedjikov, 2009) for mobile phones that support J2ME (Java 2 Micro Edition) and WAP (Wireless Application Protocol) to integrate with traditional learning and testing processes. The system was developed on the Java-based client-server communication model architecture, and comprised of six functions: Tests, Statistics, Calendar, Config, Teachers and Help.

A system to complement the structural programming course in Universiti Teknologi Petronas (UTP) was proposed by (Wendeson et al., 2010). It contains five modules: Lecture materials, Assignments, Academic information, Discussion and Quizzes. These modules were developed on a mobile web form based on the Microsoft.NET framework. Researchers used Microsoft Mobile Internet Toolkit (MMIT), which is an extension of the Microsoft.NET framework, so they can generate different mark-up languages for different devices such as HTML for PDAs, and WML for WAP phones.

The system of Mobile Learning Quiz (MLQ) was proposed by (Niazi and Mahmoud, 2008) to facilitate instructors in creating quizzes. Once the quizzes were created, they were available in many formats including XML, HTML, XHTML, WML, Java ME Midlet and Blackberry API code for supported devices. In addition, students can view their grades, instructor's feedback, messages, notifications, and access quizzes both online via a WAP browser and offline via a Blackberry or Java-enabled device. The researchers also planned to design and develop tools which instructors can use for creating course materials (e.g., lecture slides, notes) that can be automatically regenerated for different mobile devices.

El-Sofany and El-Seoud (2009) developed the Wireless Course Management System (WCMS) as a WAP application. Users including the instructor, students and the system administrator are connected to the system through a WiFi (Wireless Fidelity) network. The instructor uses the system to create the syllabus, schedule, assignments, laboratories, course resources, tutorials, tests/quizzes, grades, useful links, SMS and e-mail. Students can log into the system to view a course syllabus and do tests/quizzes. The system administrator performs the tasks of user management, course management and course database maintenance on the Web server.

Web 2.0 techniques such as Wiki, Really Simple Syndication (RSS 2.0) and Asynchronous JavaScript and XML (Ajax) were used by (Tai and Yang, 2008) to create a mobile collaborative learning platform. The system comprises teaching material, a discussion system and a Wiki system with editing, rating and category extensions. Also, learners can browse learning material and make comments with web browsers in Windows mobile devices.

2.1.3 Multimedia mobile content authoring tools

Prototypes of two mobile multimedia learning materials projects (Trifonova and Ronchetti, 2003) showed how to create multimedia content with macromedia Flash for mobile learning on a PDA, including a mobile local history tour project to support informal learning for adults, and a first-year Java programming learning objects project at London Metropolitan University. The mobile local history tour named 'Lost Worlds of Somers Town' provided multimedia information and a map with highlighted routes for users to follow. Users can learn, explore and compare the past with the present of eight historical places in Camden, London. In the learning objects for Java programming project, the learning content was adapted from a small, self-contained unit of interactive learning material, which was used to explain learning objectives suitable for students to finish within 5-10 minutes. Researchers had to reconsider the screen layout, navigation and user-control buttons on the PDA. They replaced the explanatory text with audio, reduced animations' scale and adapted some interactive user control elements of the original content to retain original pedagogy and fit within the PDA's full screen dimension of 320 x 240 pixels.

An Interactive Mobile Learning System based on PDAs using Windows Mobile 5 operating system was proposed by (Sitthiworachart, 2007) to support mobile learners in the Human-computer interaction (HCI) course at King Mongkut's Institute of Technology North Bangkok, Thailand. The system comprised three modules including multilingual computer assisted instruction (CAI), an interactive web board and a class alert system. Firstly, the multilingual CAI was designed similar to playing a movie and students can use functions such as play, pause, stop, fast forward and rewind to study the content on their PDAs. In addition, students can listen to and read in three supported languages, Thai, English and Indonesian. Secondly, the interactive web board was built to support communication among students, their classmates and the instructor in the HCI class by allowing students to choose between being alerted to responding to issues/questions through SMS or e-mail. Lastly, the class alert system, which was a simple calendar, can be used by instructors and students to preset their class activities and send alerts such as assignment deadlines, class cancellations, and class information via SMS and/or email.

An educational adventure game authoring environment called <e-Adventure >was developed by (Lavín-Mera et al., 2009) for instructors to create educational games of low complexity and low cost. The system provided a game editor, a game engine written in J2ME and a mobile device profile database described as a Wireless Universal Resource File (WURFL) (Passani and Trasatti, 2004). Instructors can create a simple game and export it to a selected mobile platform.

A Mobile E-learning Platform for All (MEPA) (Bai, 2010) based on Apple's iPhone, PhoneGap and Google Apps Engine was developed allowing non-technical instructors to develop mobile learning content by using web technologies like HTML, CSS, and JavaScript. Instructors can use web authoring software such as Microsoft Word, Macromedia Dreamweaver or Microsoft FrontPage to produce instructional content on their desktop and transfer the content into the cloud to be accessed from students' mobile phones. This platform used a hybrid application development approach through PhoneGap, an Open Source Mobile Framework that supports 6 platforms (iOS, Android, Blackberry, Palm, Window Mobile, and Symbian). In addition, the system provided mobile themes or styles for instructors to adapt the content such as announcements, weekly schedules, lecture notes, and readings online, which were usually posted on the Learning Management System, to mobile phones. Furthermore, the instructors can use an authoring environment to create quizzes in structured text format (CSV or XML), and they can track students' scores through reports in the server. My attention shifted to the delivery of online contents to students who access them through their mobile devices. This become the second part of this literature in Section 2.2.

2.2 Mobile technologies for collecting students' feedback

Besides the use of mobile phones to deliver learning contents in mobile learning systems, they have been used to gather students' feedback while a lecturer is teaching. Students' feedback is important for teaching and learning process because it helps the lecturer know the students' learning behaviour and understanding in lecture. Feedback needs to be taken in order to improve teaching (Poulos and Mahony, 2008) and develop students (Carless, 2006; Sadler, 2010). However, students often act as passive viewers in the classroom and expect the lecturer to feed them with information. Students rarely have the confidence to ask questions, ask lecturers to slow down, repeat something or explain a topic further when they are unsure of the subject matter, especially for international students who come from different backgrounds, cultures and experiences with different teaching methods.

Generally, students raise their hands to ask or answer questions, but it is not suitable for everyone particularly those who are not very confident. Seating arrangement (Roth et al., 1999) and academic achievements (Shachar and Sharan, 1994) also determine students' opportunities to interact with their teacher. In addition, it is even harder for students to interact with their lecturer in a large lecture class. A lecturer is often unaware of how her/his students are grasping concepts in a class because it is hard to know how much students are engaged in the lecture. On the other hand, if every student wants to ask a question, it may be time-consuming and the lecturer won't be able to finish the prepared material in the allocated time. Therefore, students often seem bored due to the impersonal and one-sided environment of large lecture class (Hensley and Oakley, 1998; Gehringer, 2012).

During the past ten years, students' feedback have been collected through different kinds of devices, such as clickers and mobile phones. A study showed that the use of clickers can better engage students (Beekes, 2006) and increases the students' participation more than raising hands (Denker, 2013). However, in addition to cost of providing a device for each participant (Bär et al., 2005; Beekes, 2006), clickers have the disadvantage of students losing them, breaking them or forgetting to bring them to class (Denker, 2013).

To solve the problems of clicker cost and limited information provided by clicker (Teevan et al., 2012), mobile phones came as an alternative solution because they were popular among students and students can input text with them. Furthermore, almost all students (96.8%) owned a mobile phone that they normally had with them during classes (Scornavacca et al., 2009), so there was no extra cost for the students or the university. Mobile phones can be used to send feedback using SMS, social media and mobile applications.

2.2.1 Short Message Service (SMS)

Short message service (SMS) is a component of the Global System for Mobile Communications (GSM) series of standards in 1985. SMS is a means of sending short text messages not longer than 160 characters (including spaces) between mobile phone devices. The idea of collecting feedback through SMS in the education system was proposed in (Leong et al., 2012). In this work, researchers wanted to improve the delivery of the lesson by finding out students opinions. They developed three models: the base model, the correct model and the sentiment model. The base model analyses the whole SMS text corpus without considering spelling errors. The corrected model adjusts spelling errors in SMS and classifies it under the same concept. For instance, the word 'slp' will be classified under the concepts sleep and sleepy, while 'explain' and 'explanation' are both classified under the concept explanation. Lastly, the sentiment model performs sentiment mining on corrected SMS to find interestingness and classifies the concepts into true and false. These models explore the potential application of sentiment mining for analysing short message service (SMS) texts in teaching evaluation.

Although the system had many benefits from taking feedback via SMS, it inherited limitations from SMS such as the maximum number of characters and the incompleteness of message, for example, a respondent may send a single alphabetical letter instead of a complete message. Also, words can have different meanings and can be positive and negative according to the students' context. In addition, one aspect that the authors (Leong et al., 2012) did not take into consideration was the cost of the SMS texting service. Also, the authors only tested the students' feedback at the end of semester and consequently the results can only used to improve teaching in the next semester. It would be more effective if the feedback was obtained during the semester to ensure that lecturers can adjust their teaching in time and students get the best possible learning experience. It was suggested in (Leong et al., 2012) that including a trend analysis (Lent et al., 1997), which requires the stamping of date and time on SMS texts, would help them improve their teaching over time.

2.2.2 Social media

Social media is the most common tool for enabling backchannel discussion, and some institutions have had success using social networking tools such as Twitter as a learning environment (Aagard et al., 2010). It is estimated that over 470 universities worldwide are using social networks such as Facebook and Twitter to communicate with students (Novak and Cowling, 2011). Twitter's advantage in education is that students are already familiar with the tool, so there is no need for training (Novak and Cowling, 2011). Another advantage is that Twitter solved the issue of the SMS cost because Twitter is a free service which students can access from their own Smartphones using their university wireless network without extra cost.

In addition, a microblogging style post in Twitter can be seen as an advantage to students as they need to express their thought concisely within 140 characters. This makes them reflect on their posts to create a sentence as meaningful as possible, and it could lead to a deeper understanding of arguments posted by them and others (Kuhn and Goh, 2005).

However, Twitter has some disadvantages. Firstly, in using Twitter, it requires everyone to sign up a Twitter account. So, in a large classroom, it would be a cumbersome process if instructors need to record which student is connected to which account (Aagard et al., 2010). Secondly, the tweets are shown chronologically, so users have to read from the beginning to understand what is going on, therefore it is a time consuming process (Gehringer, 2012). Lastly, it would be difficult to maintain privacy or restriction of the conversation because anyone in the world with a Twitter account can eavesdrop on a classroom discussion (Aagard et al., 2010).

2.2.3 Mobile applications

Mobile devices such as smartphones and tablets provide a good solution for backchannel systems because of their ability to run applications and their widespread adoption in classrooms (White and Turner, 2011). Following are some of the backchannel systems that are available on mobile devices, as represented by Mobile Lecture Interaction (Costa et al., 2008), Hotseat (Aagard et al., 2010), Backstage (Gehlen-Baum et al.), ActiveClass (Ratto et al., 2003) and ClasCommons (Du et al., 2012).

A system known as the Mobile Lecture Interaction (MLI) (Costa et al., 2008) was developed at the University of Oulu, Finland. Students can pose questions from their mobile phones to the lecturer and other students can support them by voting for their questions. On students' phones were Java applications connected to a web server which sent the posed questions to the lecturer on their PC. The lecturer can subsequently answer the posed questions as she/he felt the need. When the system was tested, lecture interaction improved significantly.

Hotseat is a mobile backchannel system that supports microblogging style discussions both in and out of a classroom. During a class, the lecturer can use the system to provide questions and comments to the students, who can use their mobile devices to give feedback to the lecturer, read, vote, and comment on posts from other students. Hotseat has a user-friendly interface to enable users to quickly read posts, vote, answer relevant posts, and mark favourite posts for later reviews. Each discussion is classified based on posts that are “fresh” (most recent), “hot” (most popular), and “deep” (most discussed).

Backstage is another mobile backchannel system that supports different forms of communication between students via microblogging style messages, social evaluation, and ranking of messages by the audience. Backstage emphasises anonymous and pseudonymous forms of inter-personal communication in large lectures.

ActiveClass is a simple client-server application designed to enhance participation in large classroom settings via small mobile wireless devices. Students in a class can use PDAs or low-cost mobile devices with wireless connections to anonymously ask questions, answer polls related to the questions, and give the lecturer feedback on the class through a mobile web interface. The lecturer and all students can see lists of the questions and poll results. Furthermore, students can vote on questions which they find interesting to encourage the lecturer to answer those questions.

ClasCommons is a public backchannel for building community feelings among students in university courses. Students can post messages to the system through any device with web browsing capability such as web-enabled mobile phones and laptops. Then, the messages will be displayed in real time and in chronological order on a public display in the front of the classroom, which is viewable to the students and the lecturer. Students can respond to the posted messages via the client interface and vote up/down individual messages through 'likes' or 'dislikes'.

Common to these systems is the focus on the feedback management. In other words, the user interface has been specifically designed to make it easy for students to input feedback and read others' posts. However, not much attention has been given to help the feedback consumers, - the lecturers - to easily grasp the aggregated feedback from the crowd and respond to the most important concerns students share in common. In short, current mobile backchannel systems are not capable of providing lecturers immediate and meaningful responses as we present the common limitations of these systems as follows:

1. Lack of support for lecturers: Due to a limited teaching time to a large audience, it is difficult for the lecturers to process and respond to a large amount of students' feedback in real time (Cetintas et al., 2011). For example, in Hotseat, lecturer needed to use two teaching assistants on laptops to answer students' questions as they came in, and every 15 minutes or so the instructor paused his lecture for a few minutes and took questions gathered by the teaching assistants. Similarly, results of the ClasCommons experiments

showed that lecturers had difficulties keeping their preferred teaching styles and lecture paces as they were distracted by following students' messages on the public display. In short, existing backchannel systems do not much support to providing lecturers immediate and meaningful responses.

2. Scattered and overloaded information: An important problem of microblogging supported backchannel systems is that the overall number of posts a teacher receives from the students can be overwhelming in a limited time (Cetintas et al., 2011). Most of the systems sort posts according to most recent or a histogram of class response for each question (Dufresne et al., 1996; Aagard et al., 2010; Cetintas et al., 2011; Du et al., 2012; Gehlen-Baum et al., 2012); however, this method is difficult for lecturers to quickly summarise and evaluate the overall attitude of the students in the class at any given moment and adapt their teaching strategies spontaneously to respond to students feedback.

3. Lack of support for sentiment and emotion analysis: Even though some of the current backchannel systems provide emoticons for students to select, the systems do not process those emotions embedded text and display to a lecturer in a meaningful way and in real time (Bergstrom et al., 2011; Du et al., 2012).

2.3 Emotions in learning

In addition, students' learning often involves emotions. The literature on emotions and learning has pointed out a number of human feelings related with the learning context and academic achievement, such as anger (Vohs et al., 2007; Pekrun et al., 2002), boredom (Artino, 2012; Noteborn et al., 2012), desire (Cleveland-Innes and Campbell, 2012), enjoyment (Artino, 2010; Zembylas, 2008), happiness (White, 2013), pride (Regan et al., 2012) and yearning (Cleveland-Innes and Campbell, 2012). Also, learning how to control emotion states makes students become a better learner (Falout et al., 2009).

However, in the past, emotion and cognition were considered as polar opposites. Early philosophers sought to remove emotions from the process of learning (Oregan, 2003). Over the last

15 years, the interest in the emotional aspects of learning in traditional classrooms has grown considerably. Educational researchers have recognised emotions as being an important factor to consider in learning and teaching (Pekrun and Linnenbrink-Garcia, 2012). Many studies have found that learners' emotions have an effect on motivation, self-regulation and academic achievement (Schutz and DeCuir, 2002; Sylwester, 1994), learners' study mode (Abdous and Yen, 2010; Artino, 2010) and instructional design (Gläser-Zikuda et al., 2005; Meyer and Turner, 2002).

To understand the association of students' emotion states and learning, six possible emotion axes that may arise in student learning were proposed in (Kort and Reilly, 2002), as shown in Figure 2.1, and a model of affective states and the learning cycle in science, technology, engineering and mathematics (STEM) education, as shown in Figure 2.2. The Kort and Reilly's learning cycle model suggests that students begin in quadrant I or II with curiosity and fascination about a new topic of interest in quadrant I or motivation to reduce confusion in quadrant II. As students progress and experience challenges in learning, they move down into the quadrant III, where they may feel frustrated, try to eliminate misconceptions and reinforce their knowledge with a sense of making progress and move to quadrant IV. With a new idea, students get back into the quadrant I and the cycle continues as students are presented with other challenges or new knowledge in the topic.

In summary, these studies have confirmed the importance of emotions in classroom teaching and learning, and recognising and reacting to students' emotions is crucial for effective learning and teaching in the classroom.

Axis	**-1. 0**	**-0. 5**		**0**	**+0. 5**	**+1. 0**
Anxiety-Confidence	Anxiety	Worry	Discomfort	Comfort	Hopefulness	Confidence
Ennui-Fascination	Ennui	Boredom	Indifference	Interest	Curiosity	Fascination
Frustration-Euphoria	Frustration	Puzzlement	Confusion	Insight	Enlightenment	Euphoria
Dispirited-Enthusiasm	Dispirited	Disappointed	Dissatisfied	Satisfied	Thrilled	Enthusiasm
Terror-Excitement	Terror	Dread	Apprehension	Calm	Anticipatory	Excitement
Humiliated-Proud	Humiliated	Embarrassed	Self-conscious	Pleased	Satisfied	Proud

Figure 2.1: Emotion sets possibly relevant to learning (Kort and Reilly, 2002)

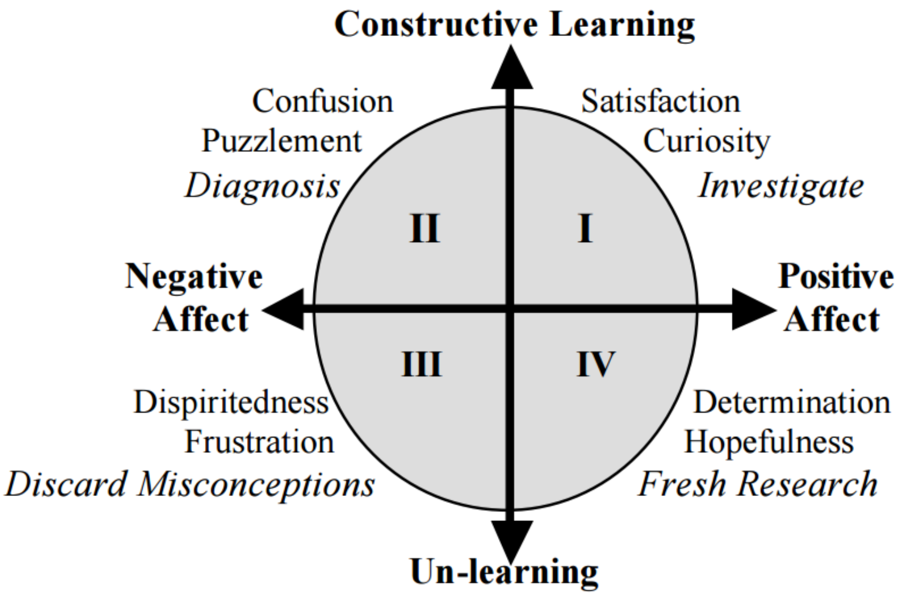

Figure 2.2: Four Quadrant model relating phases of learning to emotions (Kort and Reilly, 2002)

2.4 Emotion recognition

In small classrooms, teachers can detect students' emotions from their facial expressions or response to questions, such as "do you understand?" or "do you follow?", and change their teaching strategies accordingly (Kung-Keat and Ng, 2016). According to Visschedijk et al. (2013), human emotions can be recognised through the combination of posture,facial expression, and tone of voice, also there is a strong connection between emotions, physiological signs (e.g., pulse, blood pressure) and learning behaviour Chen and Lee (2011). However, detecting students' emotions in a large classroom and an online learning system is more difficult (Reyes et al., 2012; Binali et al., 2009).

In large classrooms, although backchannel systems have been implemented to gauge students' responses, the existing systems have no support for emotion detection yet, as we mentioned in Section 2.2.3. In an online learning context, this problem is even worse because students and lectures can't see each other, so they have to rely only on text communication. Online students need to intentionally establish an online presence by engaging in peer discussions and other online activities, and lecturers need to be able to pick up students' emotions hidden in communication texts (Hara, 2000), as reported by Oregan (2003), online students also experience several emotions, such as frustration, fear, anxiety, pride and embarrassment, while they are learning.

Some of the most suitable emotion recognition methods for educational context are using facial expression and texts analysis. Emotion recognition from facial expression is an affordable and mature method because it requires only a video camera, which is often embedded in most laptops, tablets and mobile phones, and research in this area has reached a high level of precision and maturity (Fasel and Luettin, 2003; Bartlett et al., 2006). Hwang and Yang (2009) employed image processing technologies to automatically recognise students' affective states, such as drowsy, inattentive and confused, from their facial images, so teachers can efficiently manage students' behaviours in distance learning.

For text analysis, with the rising of integrating social networking platforms like Twitter and Facebook in higher education (Mazer et al., 2007; Ortigosa et al., 2014; Maleko et al., 2013; Cloete et al., 2009; Roblyer et al., 2010), texts and emoticons posted by students in these sites are also a rich source of information to recognise students' emotions (Hancock et al., 2007). A short text used in all social networking websites is known as "microblog". It allows users to quickly send short text updates, usually less than 140 characters, to share and exchange their ideas with a specific group, individual or public (Kaplan and Haenlein, 2011; Java et al., 2007; Ebner and Schiefner, 2008). Also, emoticons are widely used to express people's emotions in microblog posts (Xue et al., 2014; Wen and Wan, 2014).

In conclusion, we believe that microblogs are more suitable to convey students' emotions in both large classrooms and online learning systems through our cross-platform application. Our decision is based on some drawbacks of facial expression analysis from the literature (Hwang and Yang, 2009; Eveland et al., 2003) and from the fact that students can't look at their mobile phones or laptops all the time while they are studying in a classroom.

Researchers have developed a technique commonly known as sentiment analysis or opinion mining, to automatically detect and classify emotions and opinions towards individuals, organizations, products, services or events from a large amounts of online texts, especially microblogs. In the next section, we present some of the sentiment analysis techniques and a detailed overview of previous studies that applied sentiment analysis with microblogs.

2.5 Sentiment analysis

Sentiment analysis is within the area of natural language processing (NLP) and is generally defined by Pang and Lee (2008) as the computational treatment of opinions, feelings, emotions, and subjectivity in texts. The sentiment found within comments, feedback or critiques is usually categorised either into two categories: positive and negative; or into an n-point scale, e.g., very good, good, satisfactory, bad, very bad (Prabowo and Thelwall, 2009). In addition, a word can have both positive and negative meanings depending on the sentence it is put in. For example, the word small can be negative if describing a hotel room and positive if describing a mobile phone. Also, the same product may get different opinion from users, so the sentiment analysis should also focus on understanding rationale and reasoning behind those opinions with respect to a specific features of product (Rahayu et al., 2010).

Sentiment analysis has been used to extract the sentiment polarity (positive, neutral, negative) and emotions of texts in various genres including news headlines (Strapparava and Mihalcea, 2007), marketing (Chamlertwat et al., 2012), politics (Mullen and Malouf, 2006) and movie reviews (Pang and Lee, 2008). Recently, sentiment analysis has been applied in the educational context including e-learning (Rodriguez et al.; Ortigosa et al., 2014) and students' learning diaries (Munezero et al., 2013). Tian et al. (2009) created patterns to find what Chinese words are associated more with emotions in an e-learning system. Analysing text can help the lecturer understand more about her/his students, reduce emotional distance between the lecturer and the students, and improve the quality of teaching and learning.

Regarding the techniques used for sentiment analysis, two main approaches are considered: machine learning methods and lexicon-based approach (Pang and Lee, 2008). Machine learning methods often produce only a binary result (positive or negative) from batch processing using a supervised classifier with a large domain-specific set of labelled training set, so that the classifier can distinguish between positive and negative patterns of messages.

Some of the most popular machine learning algorithms for sentiment analysis are Support Vector Machines (Pang et al., 2002; Dave et al., 2003; Gamon, 2004; Matsumoto et al., 2005; Airoldi et al., 2004), Naive Bayes (Wiebe et al., 1999; Yu and Hatzivassiloglou, 2003; Melville et al., 2009), and Maximum Entropy-based classifiers (Nigam et al., 1999; Pang et al., 2002). According to Pang et al. (2002), researchers applied these three algorithms with movie review data and found that Support Vector Machines appeared to be more effective than Naive Bayes and Maximum Entropy. Also, Groot (2012) concluded that Support Vector Machines achieved a better performance than the unsupervised algorithms when using it for opinion analysis on Twitter, but Go et al. (2009) concluded that Naive Bayes, Maximum Entropy, and Support Vector Machines have similar accuracy for classifying sentiment in Twitter. Also, Altrabsheh et al. (2013) suggested that the combination of these techniques will produce better accuracy. Some of the previous studies that applied supervised machine learning techniques with microblogs are presented in the next sub-section.

2.5.1 Studies using supervised machine learning approaches

Go et al. (2009) studied a problem of classifying tweets as either positive or negative by using different machine learning classifiers and feature extractors. The machine learning classifiers are Naive Bayes, Maximum Entropy, and Support Vector Machines. In order to train the supervised machine learning classifier, they downloaded a large amount of tweets via the Twitter API and used a distant supervision technique to label downloaded tweets by using emoticons in the tweets as noisy labels. Then, they removed emoticons, as they were used as labels, replaced Twitter usernames and URLs with predefined tokens and replaced adjacent repeated letters with two letters. Furthermore, retweets, duplicate tweets and any tweet containing both positive and negative emoticons were removed. A total number of tweets for final training data is 1,600,000 tweets: 800,000 for each class. A test data was downloaded via the Twitter API and manually labelled. It consists of 182 positive tweets and 177 negative tweets. For feature extractors, (Go et al., 2009) used unigrams, bigrams, unigrams and bigrams, and unigrams with part-of-speech tags. In their experiments, they explored the usage of different feature extractors with the Naive Bayes, Maximum Entropy and Support Vector Machine classification methods and compared accuracy. The best result is 82.9% accuracy using SVM with unigrams. A combination of unigrams and bigrams resulted in an increase of the Naive Bayes and Maximum Entropy performance, but a decrease in the case of Support Vector Machine. They also reported that adding negation as an explicit feature with unigrams, using part-of-speech tags and using only bigrams as features did not improve classification performance.

Similarly, Pak and Paroubek (2010) collected a large amount of tweets via the Twitter API and used positive and negative emoticons in the tweets to form two types of training dataset: positive and negative sentiments. They also retrieved tweets from 44 Twitter accounts of popular newspapers and magazines, such as "New York Times", "Washington Posts" etc, to create another training dataset of objective tweets. Then, they filtered out URL links, Twitter user names and special words (such as "RT"), emoticons and stopwords from all tweets in the training datasets and tokenised on whitespace and punctuation. A negation words, such as "no" and "not", were attached to a word which precedes it or follows it. They experimented three sentiment classifiers: the multinomial Naive Bayes, Support Vector Machines and Conditional Random Fields with unigrams, bigrams, and trigrams and found the Naive Bayes classifier with bigrams yielded the best performance.

Barbosa and Feng (2010) proposed a 2-step sentiment detection framework to classify a tweet into one of the three sentiment classes: positive, neutral or negative. The first step is subjectivity classification and the second step is polarity classification. Instead of using n-grams as features, both classifiers perform prediction based on two sets of features: meta information about the words on tweets (e.g., part-of-speech tags, prior subjectivity and polarity of words) and tweet syntax features (e.g., retweet, hashtag, URL, punctuation, emoticons and upper case token). In the evaluation, they get the best results using a Support Vector Machine classifier and the two sets of features for both steps by achieving 81.9% accuracy for the subjectivity detection step and 81.3% accuracy for the polarity detection step, while a unigrams baseline achieved only 72.4% and 79.1%, respectively.

Bermingham and Smeaton (2010) evaluated sentiment classification accuracy among 4 short-form textual domains including microblogs, blogs, microreviews and movie reviews by using Support Vector Machine and Multinomial Naive Bayes with features such as unigrams, bigrams, trigrams, part-of-speech tags and part-of-speech n-grams. They built a manually annotated dataset of 1,410 positive, 1,040 negative and 2,597 neutral tweets from ten topics for each of five categories, entertainment, products and services, sport, current affairs and companies. They found that using the Naive Bayes classifier with unigrams outperformed Support Vector Machine on microblog and microreviews as well as achieved the best result for binary classification with 74.85% accuracy and 61.3% for ternary classification.

Davidov et al. (2010) selected 50 hashtags and 15 emoticons as noisy labels to label a Twitter dataset provided by O'Connor et al. (2010). For sentiment classification, they used a k-nearest neighbour classification algorithm with four basic feature: single word, n-grams (2-5), pattern such as high-frequency words and content words, and punctuations. In the binary classification experiments, they reported an averaged harmonic f-score of 0.86 for emoticons and 0.80 for hashtags. When compared the contributed performance of different feature types, they found that punctuation, word and pattern features provided a substantial increase for classification quality, while they received only a marginal boost with n-grams. Also, the pattern features contributed the performance more than all other features together.

Agarwal et al. (2011) acquired 11,875 manually annotated Twitter data (tweets) from a commercial source. After removing junk tweets, they used balanced subsets containing 1,709 tweets for positive, negative and neutral classes. Then, they created two new resources for preprocessing the data: an acronym dictionary and an emoticon dictionary. The acronym dictionary has 5,184 acronyms and the emoticon dictionary was created from assigning each of 170 emoticons, listed on Wikipedia, a label from the following set of labels: Extremely-positive, Extremely-negative, Positive, Negative, and Neutral. Furthermore, Twitter usernames, URLs and negations were replaced by generic placeholders and repeated characters were replaced with three characters. They also proposed a set of new 100 features, Senti-features, which can be calculated from the tweets. For example, count of number of positive adverbs, negative verbs, sum of the prior polarity scores of words with part-of-speech of adjective/adverb/verb/noun, sum of prior polarity scores of all words, presence of exclamation marks and presence of capitalized text. The prior word polarity scores were computed using the Dictionary of Affect in Language (Whissell, 1989) and WordNet (Fellbaum, 1998). In their experiments, they compared five different models based on Support Vector Machine and reported averaged 5-fold cross-validation test results. For a binary classification task, they discovered that a classifier using only Senti-features performs similar to the unigrams baseline; however, they found the best result to be 75.39% accuracy using a combination of unigrams and Senti-features. For a ternary classification task, combining the Senti-features with tree kernel, a tree representation of tweets features, yielded the best performance of 60.83%. In both binary and ternary classification, the Senti-features performed almost as well as the unigram baseline, while the tree kernel outperformed the unigram baseline and Senti-features about 4%.

To solve language-specific characteristics of the n-grams model, Aisopos et al. (2011) proposed a new document representation model, namely the n-gram graphs, for sentiment analysis of Twitter messages. The new model considers the sequence of n-grams appearance in a tweet to model a graph, with nodes correspond to specific n-grams and edges show how close they are found in the given tweet. The n-gram graphs is language independent and it can capture more information than a plain bag of n-grams. Their training data consists of 1 million tweets for each of the three classes: positive, negative and neutral, annotated with emoticons as noisy labels. To evaluate the model, they used two classification algorithms, the Naive Bayes Multinomial and the C4.5 tree classifier, with a vector model, n-grams (2-4) models and the n-grams graphs. A performance comparison among the models showed that the n-grams graphs achieved lower accuracies for the Naive Bayes Multinomial, but took substantially higher values for the C4.5 in all cases. The best model is based on using 4-gram graphs with the C4.5 tree classifier and achieved 66.77% accuracy with two classes and 50.67% in case of three classes.

Kouloumpis et al. (2011) studied a variety of linguistic features for sentiment classification of Twitter messages. For training data, they used the hashtagged data set (Petrovic et al., 2010) and the emoticon data set (Go et al., 2009) as noisy labels for tweets. A hand-annotated set of 4,000 tweets from the iSieve Corporation was used as test data. In the preprocessing step, emoticons and abbreviations (e.g., BRB, LOL) were identified and replaced by actual meaning (e.g., BRB >be right back), informal intensifiers such as all-caps (e.g., I LOVE this show!!!)

were made into lower case, character repetitions (e.g., I've got a mortgage!! happyyyyyy) were replaced by a single character and any special Twitter tokens (e.g., #hashtags, usertags, and URLs) were substituted with generic placeholders. Regarding features, after removing stop-words, they used the top 1,000 unigrams and bigrams according to their information gain measured by Chi-squared, prior polarities according to the MPQA subjectivity lexicon (Wiebe et al., 2005), count of the number of part-of-speech tags and presence or absence of microblogging features (e.g., emoticons, abbreviations, intensifiers). In their experiments, an AdaBoost.MH classifier (Schapire and Singer, 2000) outperformed Support Vector Machines and achieved the best performance of 75% accuracy for sentiment classification of the three classes: positive, negative and neutral, using a combination of n-grams, lexicon and microblogging features.

Alternatively, a lexicon-based approach consists of analysing words in the target text by using a predefined sentiment lexicon - a dictionary of words annotated with their semantic orientation (sentiment polarity and strength) - and executing a function to calculate a sentiment score for a piece of text based on the predefined scores in the lexicon. A clear advantage of this approach is that lexicons are more easily available and extensible than training sets and more robust when considering cross domain applications (Turney, 2002; Taboada et al., 2011). There are several sentiment lexicons available, such as SentiWordNet (Baccianella et al., 2010), NRC word-emotion association lexicon currently including positive and negative emotional annotations for 14,182 unique words (Mohammad and Turney, 2013), LIWC (Linguistic Inquiry and Word Count) (Tausczik and Pennebaker, 2010) and SentiStrength (Thelwall et al., 2010). Some of the previous studies that applied lexicon-based techniques with microblogs are presented in the next sub-section.

2.5.2 Studies using lexicon-based approaches

In a political context, O'Connor et al. (2010) used the MPQA sentiment lexicon Wiebe et al. (2005) to measure public opinion about Barack Obama from 1 billion Twitter messages posted over the years 2008 and 2009. They calculated day-to-day sentiment scores by counting positive and negative words in tweets according to the sentiment lexicon. Although this is a relatively simple approach, they found a significant correlation between the calculated sentiment scores and public opinion polls published by Gallup. Similarly, Marchetti-Bowick and Chambers (2012) used tweets for political forecasting; however, they received better correlation with Gallup polls by using the lexicon-based sentiment analysis with distant supervision.

Thelwall et al. (2010) proposed SentiStrength, a lexicon-based algorithm which assigns a polarity (positive/negative) and corresponding strength value between 1 to 5 (positive) and -1 to -5 (negative) to a given short informal text. SentiStrength uses a list of 298 positive and 465 negative terms, emoticons, negations and boosting words annotated with polarity and strength values to analyse texts. In performance evaluation step, the authors compared SentiStrength with several machine learning classifiers, such as Support Vector Machine, Simple Logistic Regression and Naive Bayes, on 1,041 MySpace comments, and found their method performed better in classifying negative sentiment, but not for positive sentiment. The SentiStrength algorithm was improved in Thelwall et al. (2012) by including idiom lists, negating word list, an unsupervised version of SentiStrength and new terms in their sentiment strength word list (from 693 to 2310 terms). The improved SentiStrength was compared against different machine learning algorithms on 11,790 texts from six different datasets, including BBC, Digg, MySpace, Runners World, Twitter and YouTube, and found only the Logistic Regression algorithm outperformed SentiStrength.

Zhang et al. (2011) proposed a new entity-level sentiment analysis method for Twitter data. First, the authors used a lexicon-based approach to perform entity-level sentiment analysis; however, they found that this method gave high precision, but low recall. In order to improve recall, they used information in the result of the lexicon-based method to identify additional tweets that were likely to be opinionated. Then, they trained a classifier to assign polarities to the entities in the newly identified tweets by using the labelled training sets obtained from the lexicon-based approach. Experimental results showed that their methods improved the recall and the F-score, and outperformed the baselines.

Kumar and Sebastian (2012) proposed a hybrid approach which used a corpus-based method to find the semantic orientation of adjectives and a dictionary-based method to determine the semantic orientation of verbs and adverbs. Then, the overall tweet sentiment was calculated using a linear equation with emotion intensifiers.

Maynard et al. (2012) discussed the challenges of sentiment analysis imposed by social media. Then, the authors designed a modular rule-based approach which performed shallow linguistic analysis and built on a number of linguistic subcomponents to generate the final opinion polarity and score for a given tweet.

Nielsen (2011) created a new word list by including Internet slang and obscene words with sentiment strength and added it to the existing ANEW (Affective Norms for English Words) (Bradley and Lang, 1999). The author evaluated the new word list with a data set of 1,000 tweets labelled by Alan Mislove for the Twittermood/"Pulse of a Nation" (Biever, 2010) and found a slightly better performance than the existing ANEW. Also, Minocha and Singh (2012) focused on automatically building a domain-specific sentiment lexicon by using Twitter data and the ontological structure provided by Open Directory Project.

These studies provide insights into how lexicon-based method have been used recently. After carefully considering the strengths and limitations of lexicon-based methods, we chose SentiStrength and integrated it into our backchannel system *ClasSense* because: (a) it calculates a numeric score rather than just a binary result, which is essential for our morale-graph-based interface, and (b) it is open source and has been adopted by a variety of research projects.

2.6 Sentiment visualisation

To supplement the sentiment and emotion analysis, visualisation techniques have emerged to help users get a better understanding of a large collection of social media data. We select some of the related studies and report in this section. These tools use data from Twitter, blogs, customer reviews and social networking websites and convert it to formats that users can explore and validate. Opinion Space uses an interactive map of dots to allow users to visualise and navigate through a range of online comments in many topics (Faridani et al., 2010).

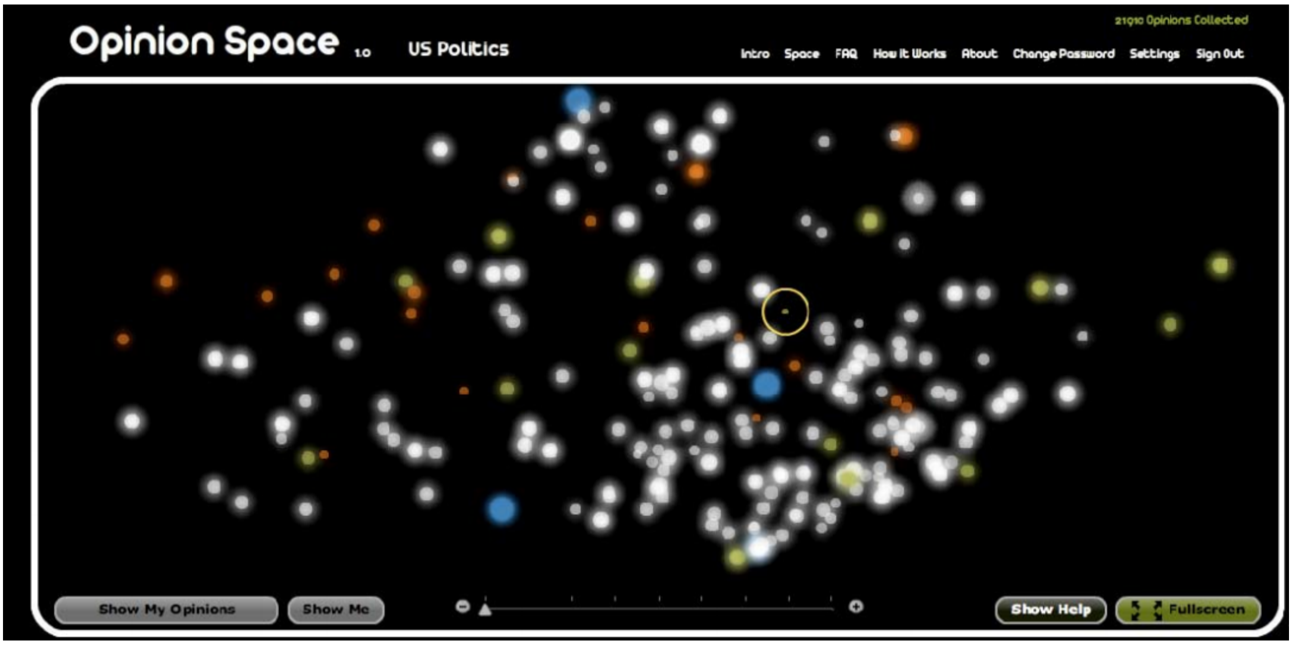

Figure 2.3: Opinion Space (Faridani et al., 2010)

Vox Civitas (Diakopoulos et al.) and TwitInfo (Marcus et al., 2011) support exploration of selected events through an interactive timeline chart that summarises the events over time from a large collection of tweets, together with the sample tweets, sentiments and locations.

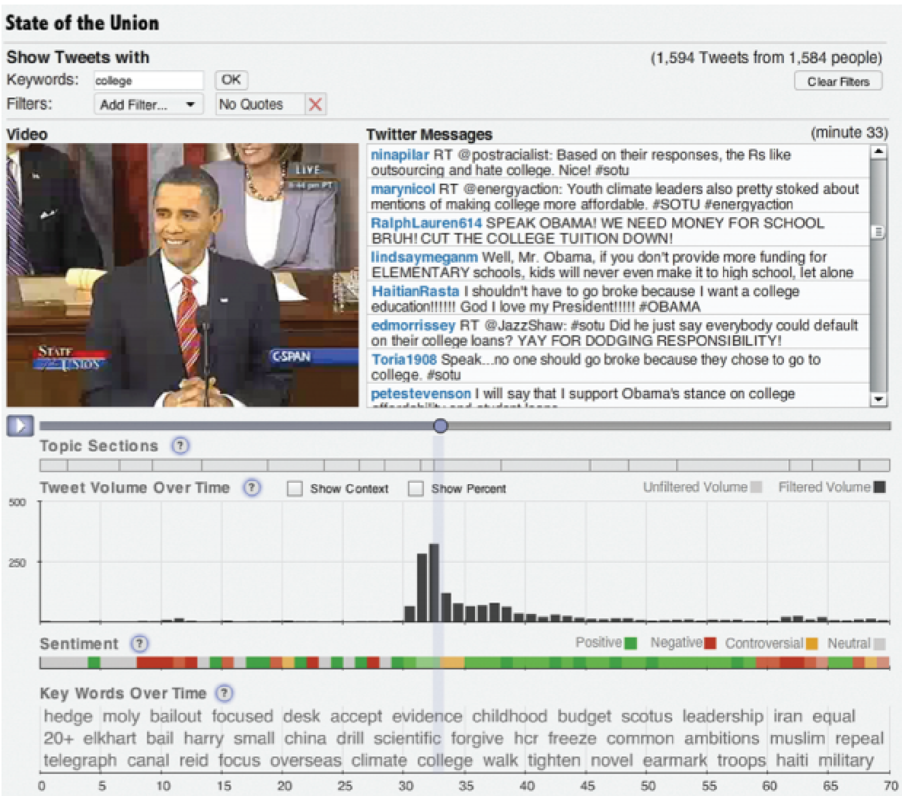

Figure 2.4: Vox Civitas (Diakopoulos et al.)

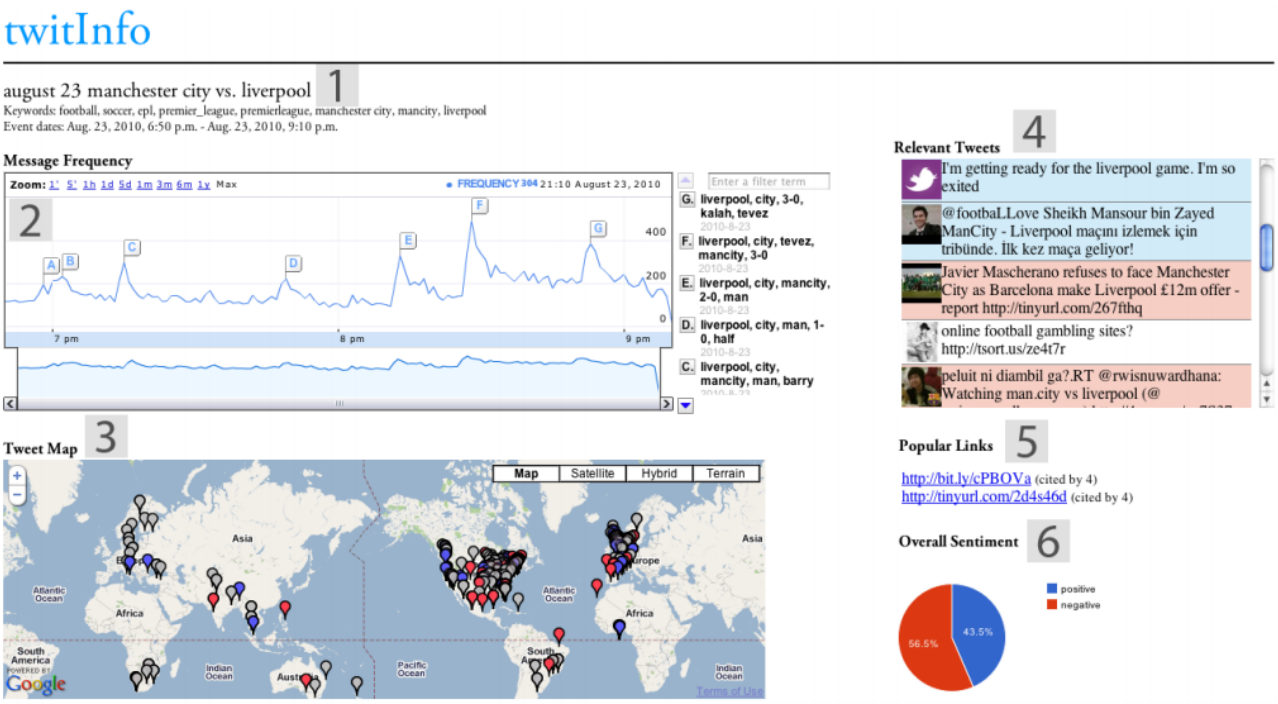

Figure 2.5: TwitInfo (Marcus et al., 2011)

In an academic context, Conference Monitor helps academic conference organisers monitor and analyse conversations on a backchannel system through a timeline chart that is associated with popular hashtags and tweets (Sopan et al.). ALAS-KA uses a line chart together with annotations of relevant events to display the variations of affective states of learners in the Khan Academy (Ruiprez-Valiente et al., 2015).

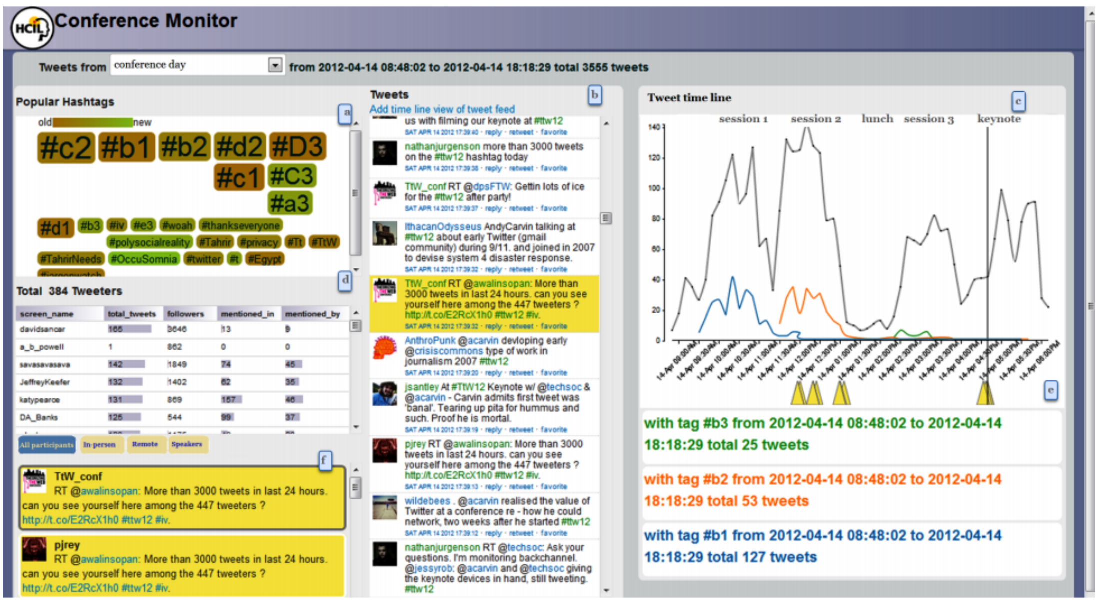

Figure 2.6: Conference Monitor (Sopan et al.)

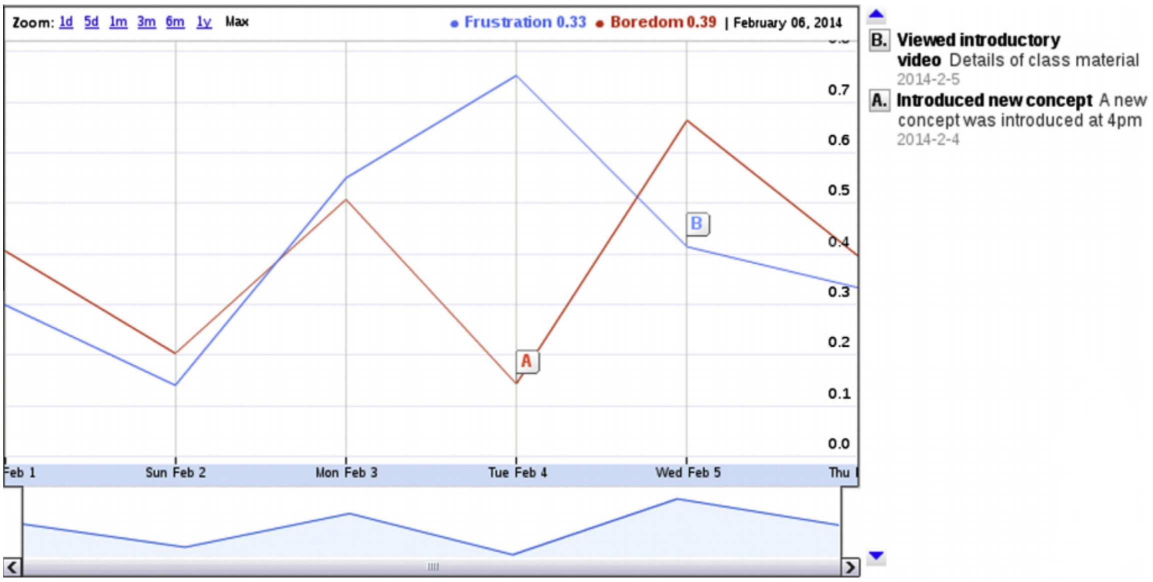

Figure 2.7: ALAS-KA (Ruiprez-Valiente et al., 2015)

Researchers integrated Synesketch (Krcadinac et al., 2013), an open source software library for emotion recognition and visualisation, with Moodle Learning Management System to allow teacher to view emotions expressed by students during their online discussion within a dedicated Moodle's discussion forum or chatroom (Krčadinac et al., 2012). Affective Tutor is developed to help instructors to determine students' feeling about the lecture pace in real-time (Hickey and Tarimo, 2014), but this work allows students to express only 3 pre-defined emotions: Engaged, Bored and Confused.

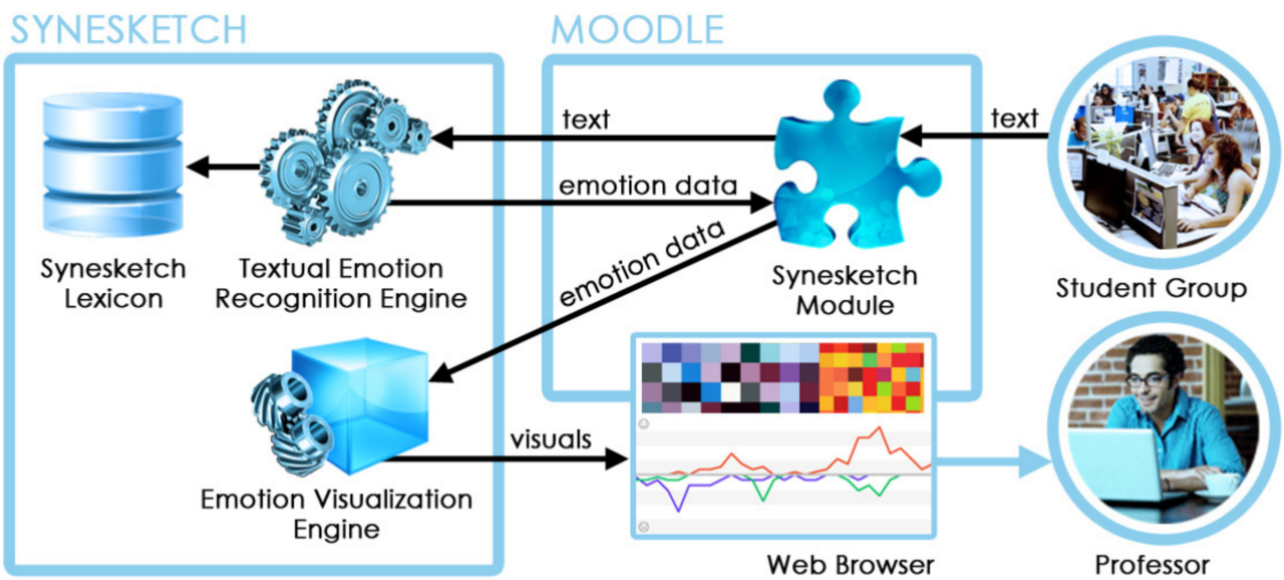

Figure 2.8: Synesketch integration with Moodle Learning Management (Krčadinac et al., 2012)

Figure 2.9: Affective Tutor (Hickey and Tarimo, 2014)

However, to the best of our knowledge, there are limited studies that applied text analytics to analyse and visualise students' affective and cognitive feedback in large traditional and online class to help improve teaching and learning (Stephens-Martinez et al., 2014).

Chapter 3

The *ClasSense* backchannel system

In this chapter, we present the design concept of the *ClasSense* system, the system architecture, the *ClasSense* student and lecturer application user interface and the techniques that we use for sentiment analysis and morale computation. Also, we explain how we customise and evaluate the SentiStrength in our context.

3.1 Using the *Seven Principles* as a guidance for the *ClasSense* backchannel system design

The *ClasSense* system design is guided by the *Seven Principles* (Chickering and Gamson, 1987), which is based on decades of research on undergraduate education and intended as guidelines for faculty members and students to improve teaching and learning. In addition, it has been widely accepted in the design and evaluation of many technology-enhanced learning environments (Junco et al., 2011; McCabe and Meuter, 2011; Suen, 2005; Koeckeritz et al., 2002). This section discusses how the principles have an influence on the design of the *ClasSense* system.

3.1.1 Encourages contacts between students and faculty

"Frequent student-faculty contact in and out of class is a most important factor in student motivation and involvement. Faculty concern helps students get through rough times and keep on working. Knowing a few faculty members well enhances students' intellectual commitment and encourages them to think about their own values and plans. " (Chickering and Ehrmann, 1996).

According to principle 1, we design the morale-graph-based user interface to help lecturers understand a trend of students' sentiments and emotions during their lecture, so they can know what students are thinking of and respond to students' feedback accordingly. This user interface also allows the lecturers to navigate a large number of posts effectively. For example, they can choose to explore posts at a particular time, which has a lot of negative messages, to see the issues that are happening in the class. Furthermore, we design the *ClasSense* student application to display lecturers' response for each student's post in real time. This feature can be used to inform students that their feedback has been acknowledged by the lecturer.

3.1.2 Develops reciprocity and cooperation among students

"Learning is enhanced when it is more like a team effort than a solo race. Good learning, like good work, is collaborative and social, not competitive and isolated. Working with others often increases involvement in learning. Sharing ones ideas and responding to others' improves thinking and deepens understanding." (Chickering and Ehrmann, 1996).

According to principle 2, the *ClasSense* backchannel system is designed with peer rating feature, which allows students to view, comment on, and vote posts made by their peers by using the "thumb up" and "thumb down" icons. The peer rating feature helps students collaborate, share ideas, and respond to each other's ideas. This collaboration would then lead to a deeper level of understanding for all of the students.

3.1.3 Uses active learning techniques

"Learning is not a spectator sport. Students do not learn much just sitting in classes listening to teachers, memorizing prepackaged assignments, and spitting out answers. They must talk about what they are learning, write reflectively about it, relate it to past experiences, and apply it to their daily lives. They must make what they learn part of themselves." (Chickering and Ehrmann, 1996).

According to principle 3, we integrate microblogging feature into the *ClasSense* backchannel system to promote active learning. Students can use microblogs to exchange their thoughts, express their opinions and emotions and share a lecture content with their peers and the lecturer. So, they have more chances to reflect on the lecture content by reading, writing and discussing about what they are learning.

3.1.4 Gives prompt feedback

"Knowing what you know and don't know focuses your learning. In getting started, students need help in assessing their existing knowledge and competence. Then, in classes, students need frequent opportunities to perform and receive feedback on their performance. At various points during college, and at its end, students need chances to reflect on what they have learned, what they still need to know, and how they might assess themselves." (Chickering and Ehrmann, 1996).

According to principle 4, the *ClasSense* backchannel system helps lecturers respond to students' queries and problems quicker by using the morale-graph-based user interface in both traditional or online lecture and during or after a lecture to provide opportunities for students to ask questions.

3.1.5 Emphasises time on task

"Time plus energy equals learning. Learning to use one's time well is critical for students and professionals alike. Allocating realistic amounts of time means effective learning for students and effective teaching for faculty." (Chickering and Ehrmann, 1996).

According to principle 5, the *ClasSense* backchannel system is designed to allow lecturers to know students' feedback while they are teaching, so they can adjust their teaching in time without waiting until the end of semester to get the students' feedback. Students also save their time by asking questions through the *ClasSense* system when they have something to ask instead of waiting to the end of lecture session.

3.1.6 Communicates high expectations

"Expect more and you will get it. High expectations are important for everyone for the poorly prepared, for those unwilling to exert themselves, and for the bright and well motivated. Expecting students to perform well becomes a self-fulfilling prophecy." (Chickering and Ehrmann, 1996).

According to principle 6, a lecturer can refer to the morale graph before communicate her/his expectation with the class. For example, when the lecturer see many negative posts, the lecturer may choose to pause her/his lecture, navigate the morale graph and posts to see what students are thinking in order to communicate her/his expectation with them appropriately.

3.1.7 Respects diverse talents and ways of learning

"Many roads lead to learning. Different students bring different talents and styles to college. Brilliant students in a seminar might be all thumbs in a lab or studio; students rich in hands-on experience may not do so well with theory. Students need opportunities to show their talents and learn in ways that work for them. Then they can be pushed to learn in new ways that do not come so easily." (Chickering and Ehrmann, 1996).

According to principle 7, we integrate an anonymous posting feature in the *ClasSense* student application to allow students to post without disclosing their real name or email. This feature encourages the students with different cultural background or shy students to ask anything they want to know without being afraid of losing face or ridicule.

3.2 The *ClasSense* system architecture

The *ClasSense* system architecture includes two components: the *ClasSense* student application and the *ClasSense* lecturer application. The components are hosted in a cloud server, a virtual server which runs on a cloud computing environment (Armbrust et al., 2010). This infrastructure is suitable for our context because it allows us to easily scale up or down resource and server specification whenever and whatever we want to support a large class. Also, it is cost effective because we pay for what we use on an hourly basis.

The architecture of the *ClasSense* system is illustrated in Figure 3.1. The student application communicates with the server by using JavaScript Object Notation (JSON). After receiving students' feedback, the server-side application saves the feedback to MySQL database, analyses the feedback every minute using the SentiStrength and displays a real-time bubble graph of morale trend and top ranked posts based on a sentiment polarity at each minute through the *ClasSense* lecturer application, which is developed using PHP and JavaScript programming language and Highcharts, a JavaScript library for creating interactive charts in a web application.

3.3 The *ClasSense* student application

The user interface of the *ClasSense* student application is designed for mobile devices; however, students can use it on laptop or PC as well because we use a HTML5 hybrid approach (Amatya and Kurti, 2014), which exploits web technologies (HTML, CSS3 and JavaScript) to develop web applications that look like a mobile application and are accessible by typing a URL in any

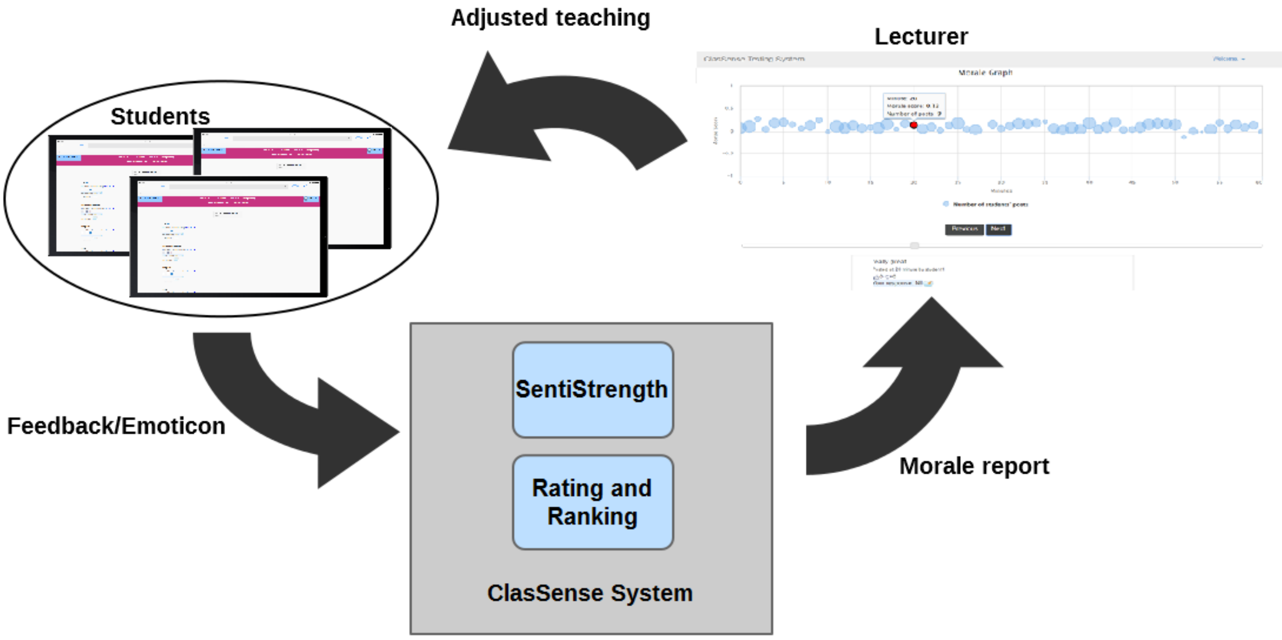

Figure 3.1: The architecture of the *ClasSense* system

browser. One of the benefits of hybrid apps is it can later be packaged with extensions that allow access to some hardware features and exported to popular native mobile platforms such as Android and iOS by using software framework such as PhoneGap and Titanium.

We used the jQuery Mobile, a HTML5 mobile application development framework, to develop the *ClasSense* student application for smartphones and tablets. The application can be accessed through students' mobile, tablet, laptop or PC browsers. Features of the application are highlighted in the following:

- Explicit display of emoticons: we clearly display emoticons in the "Create new post" page to encourage students to express their emotions in a post or comment, as shown in Figure 3.3.

- Anonymous posting: we achieve this by not using students' real name or email to encourage students to participate in a class discussion.

- Peer's rating: we achieve this by allowing students to give a rating to any post by clicking the "thumb up" or the "thumb down" icon. This rating helps inform lecturers to access the most important ones.

- Display of lecturer's response in real time: to complete a feedback loop and inform students that their feedback is acknowledged, we display the complete response from the lecturer once she/he has entered in the system; otherwise the system displays "Nil".

As shown in Figure 3.2, the *ClasSense* student application uses a microblogging user interface, which has been used in popular social networking sites (e.g. Twitter, Facebook and Google+) and recent backchannel systems like Hotseat and Backstage. Microblog allows users to quickly send short text updates, usually less than 140 characters, to share and exchange their ideas and emotions with a specific group, individual or public (Kaplan and Haenlein, 2011; Java et al., 2007; Ebner and Schiefner, 2008; Xue et al., 2014; Wen and Wan, 2014). Also, the *ClasSense* student application can be seen as an ongoing lecturer shared message board where posts are listed and accessible by students to give comments and vote whether they 'like' or 'dislike' the posts. Content of a post could be anything from a general question about the content of the lecture, positive or negative feedback about the lecture or other emotional expressions.

As shown in Figure 3.3, when a student create a new post, which is limited to 140 characters, they are encouraged to express their emotions by choosing emoticons in the "Create a new post" pop-up page to embed in a post.

Emoticons are textual representations that are inherently and immediately recognizable as faces expressing the emotions they represent. Normally, text emoticons in a post will be replaced with small smiley-faced images that correspond to each emoticon when they are selected by users; however, embedding smiley-faced images make the post not suitable for processing with SentiStrength. As a result, we choose to use the CSS Emoticons plugin, which is a simple jQuery plugin (and stylesheet) that can turn any text emoticons on the page into little smiling faces without using images. This plugin is supported by most popular browsers including FireFox, Chrome, Safari, and Opera. Also, it works on the iPhone and Android smartphone browsers. The result is great-looking emoticons that the text emoticons can actually get embedded within a post and later scored by SentiStrength.

In addition, the *ClasSense* student application can automatically display lecturer's responses once she/he has submitted to students. This feature keeps the students informed that their feedback has been acknowledged by the lecturer.

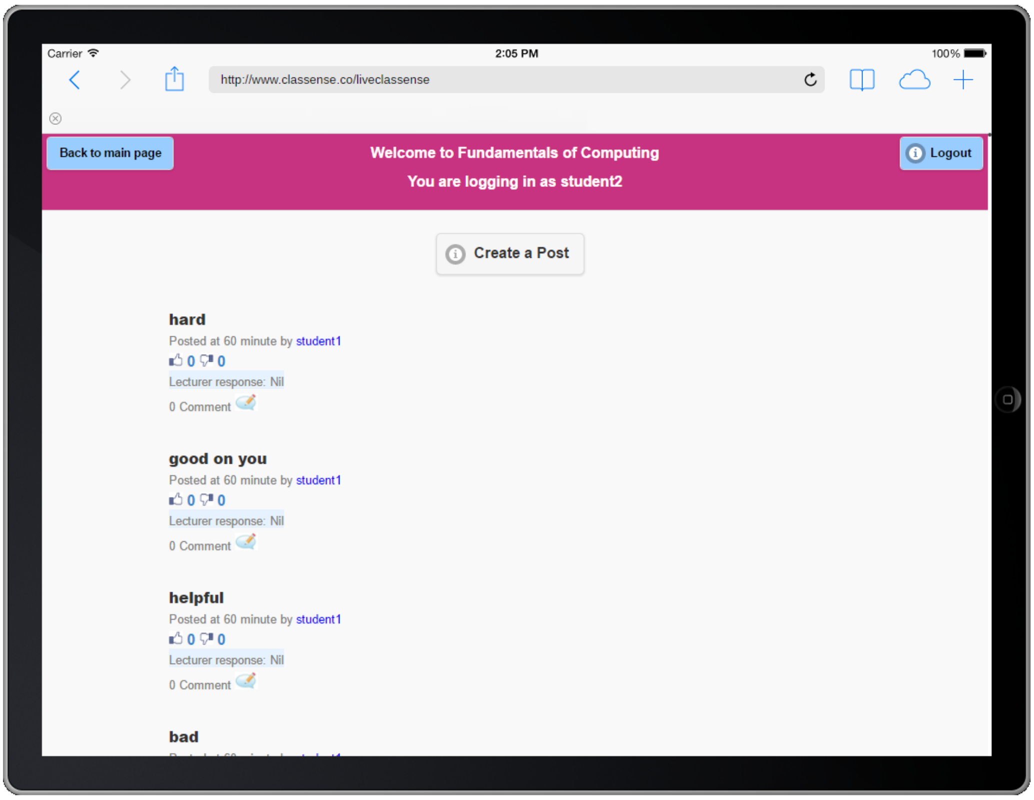

Figure 3.2: The main screen of the *ClasSense* student application

3.4 The *ClasSense* lecturer application

The *ClasSense* lecturer web application is designed and developed as a single-page application (Mesbah and van Deursen, 2007), which is a web application that fits in one page and has user interaction similar to a desktop application. The application is designed to be an intuitive and unobtrusive system that supports a lecturer viewing students' feedback in a large traditional or online lecture. Also, the application can display both online (in class) and offline (after class) morale graph, which is a trend of students' morale and associated posts, by using PHP and JavaScript programming language, JSON and the Highcharts library. By using the Highcharts library, the morale graph has become a touch-based interactive bubble graph that can be displayed in any browser and on any platform (e.g. PC, smartphone or tablet). This configuration provides the lecturer with more options to respond to students' feedback than

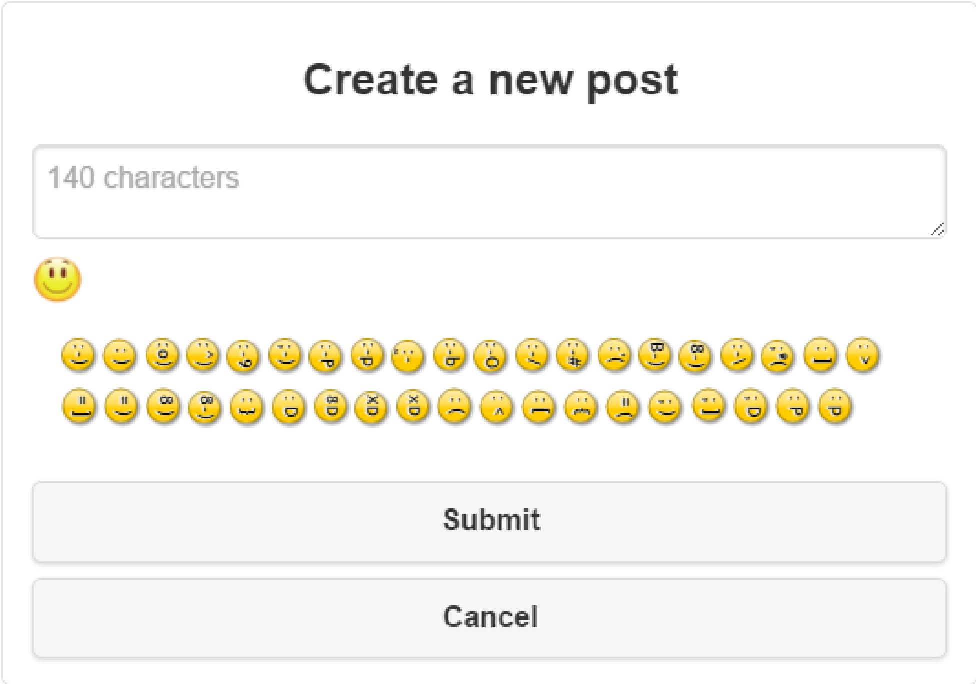

Figure 3.3: The screen for creating a new post in the *ClasSense* student application

just using a projection screen or a teacher assistant screen like most existing digital backchannel systems do.

For the *ClasSense* server application, it is developed using Python, MySQL and SentiStrength and scheduled to run every minute using Cron, a time-based job scheduler in Linux. When the server-side application starts, it selects posts that have not been processed from the database. Each post is converted to UTF-8 encoding by replacing a newline character with a plus sign. Then, the posts are passed to SentiStrength for score calculation. Once SentiStrength have finished processing a post, it assigns a morale score ranging between +5.0 (the most positive score) to -5.0 (the most negative score) to the post, which is used in calculating the y-axis (normalised morale scores) value of the graph by using the formula presented in Section 3.4.1.

3.4.1 Computation of morale scores

We have chosen SentiStrength and integrated it into the system because: (a) it calculates a numeric score rather than just a binary result, which is essential for our morale-graph-based interface, and (b) it is open source and has been adopted by a variety of research projects.

SentiStrength was designed to classify social web data, such as Tweets, MySpace and YouTube comments with a positive sentiment strength of 1 (no positive sentiment) to 5 (very strong positive sentiment) and a negative sentiment strength of -1 (no negative sentiment) to -5 (very strong negative sentiment) based on a list of 2,608 words and word stems with predefined average sentiment scores. For example, in the original SentiStrength's lexicon, "good" scores +3 and "difficult" scores 2, so the post "lecturer is good at explaining difficult topics" might get a positive strength of +3 and a negative strength of -2. If there are several sentiment words, the highest positive word score and the lowest negative word score are chosen as the post scores. In contrast, if there is no sentiment word, SentiStrength neutral sentiment scores, 1 for positive score and -1 for negative score, are set as the post scores.

Table 3.1: An example of computing morale scores using SentiStrength

Post	Word scores	[+,-] Post scores
computer make mistakes faster lol	mistakes[-2], lol[2]	[2,-2]
repetition made me lose my interest	lose[-2], interest[2]	[2,-2]
that guy in background was annoying and funny at same time	annoying[-3], funny[2]	[2,-3]
[+,-]Average morale scores of posts		[2,-2.33]
Normalised morale scores of posts		[0.4,-0.47]
Plotted morale score		[-0.47]

As shown in Table 3.1, we first used a normalised average of SentiStrength score (?) of all posts made during a particular time period to represent a class morale score *ClasSense*. The normalised morale score is calculated for positive and negative sentiments as follows:

$$MSS^p = \frac{1}{n} \times \sum_{i=1}^{n} x_i^p \tag{3.1}$$

$$MSS^n = \frac{1}{n} \times \sum_{i=1}^{n} x_i^n \tag{3.2}$$

$$NMS_t = \begin{cases} (1) \quad MSS^p & \text{if} \quad MSS^p > |MSS^n| \\ (2) \quad MSS^n & \text{if} \quad MSS^p < |MSS^n| \\ (3) \quad \begin{cases} MSS^p & \text{if} \quad NMS_{t-1} \textit{ is positive} \\ MSS^n & \text{if} \quad NMS_{t-1} \textit{ is negative} \end{cases} & \text{if} \quad MSS^p = |MSS^n| \end{cases} \tag{3.3}$$

where:

MSS^p = Mean positive sentiment score

MSS^n = Mean negative sentiment score

n = the number of posts made between time t-1 and t

x_i^p = the SentiStrength positive score of post x_i

x_i^n = the SentiStrength negative score of post x_i

NMS_t = Normalised Morale Score at time t

We then plot a morale bubble (as depicted in Figure 3.4) to represent the normalised morale score. We finally adjust the bubble size to be proportional to the number of posts and rank the associated posts based on the criteria in the order of a polarity of class morale score, number of likes, number of dislikes, number of replies, and recentness.

3.4.2 The Morale-Graph-Based user interface

In recent years, several techniques for visualising time-oriented data have been developed and used in many applications to help people understand the evolution of their data over time (Aigner et al., 2011). In this work, we use temporal visualisation of morale scores to keep a lecturer informed of students' sentiment trend during a traditional and online lecture through a morale graph associated ranked posts, which is shown in Figure 3.4.

We apply the pre-attentive processing concept (Ware, 2012) to design the visual properties (Healey and Enns, 2012) (e.g., size, colour, spatial position, mark and movement) of elements on a morale graph in order to help a lecturer process information in a short time without much cognitive effort (Spence, 2007). For example, a bubble size is used to encode the number of posts

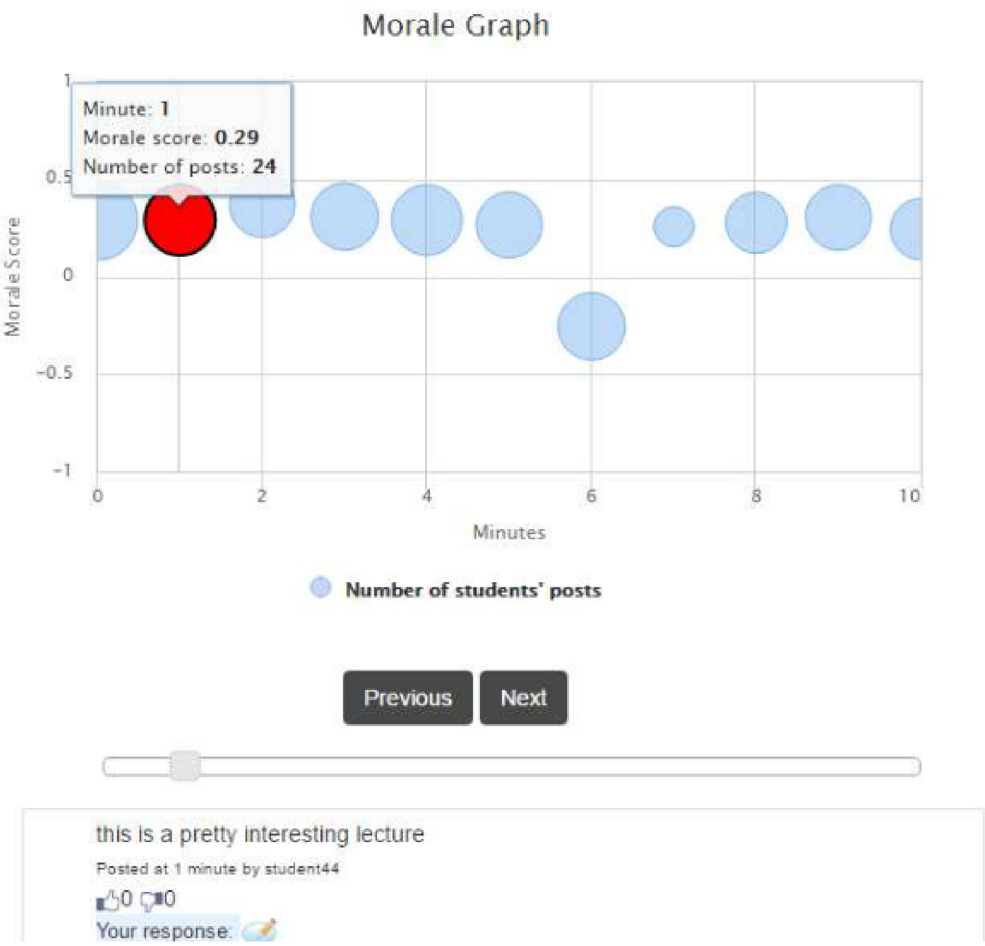

Figure 3.4: A morale graph for visualising students' sentiment

in each minute, a red colour and a vertical mark are used to distinguish a selected bubble from the rest, and a 2D orientation is used to differentiate posts in positive and negative polarities. In addition, a tooltip provides textual information about a selected bubble, including the actual moment in time, the normalised morale score, and the number of posts.

ClasSense allows a lecturer to view students' feedback both during and after a lecture. To navigate through students' feedback, a lecturer can directly click on bubbles on the morale graph or drag the knob of a horizontal timeline slider for quick selections. They can perform fine-grained minute-by-minute navigation with the "Previous" and "Next" buttons.

In addition, all user interface components are linked together to provide flexibility in accessing students' posts. When a lecturer drags the slider or presses the "Previous" and "Next" buttons, the bubble that aligns with the position value of the slider will change its colour to red and move accordingly together with the associated posts.

3.5 Pilot test

Figure 3.5 shows the morale graph and the associated posts of a pilot test in an Information Technology lecture. When the lecturer noticed a sharp morale drop at the six-minute mark of the lecture, they immediately clicked on the corresponding bubble (in red) to investigate what happened from the associated posts. From the posts that were ranked based on a negative polarity (maximum to minimum) as the morale score is negative, the lecturer discovered the most important issues students had, such as *cards and examples did not work well, distraction and interruption in the class*, etc. As a result, they opted to go through the examples again in the lecture but to write a response about the cards after the lecture.

3.6 Customisation of SentiStrength Lexicon

The SentiStrength lexicon is intended for general-purpose sentiment strength detection. However, there may be occasional or specific words that are frequently used to express sentiment in

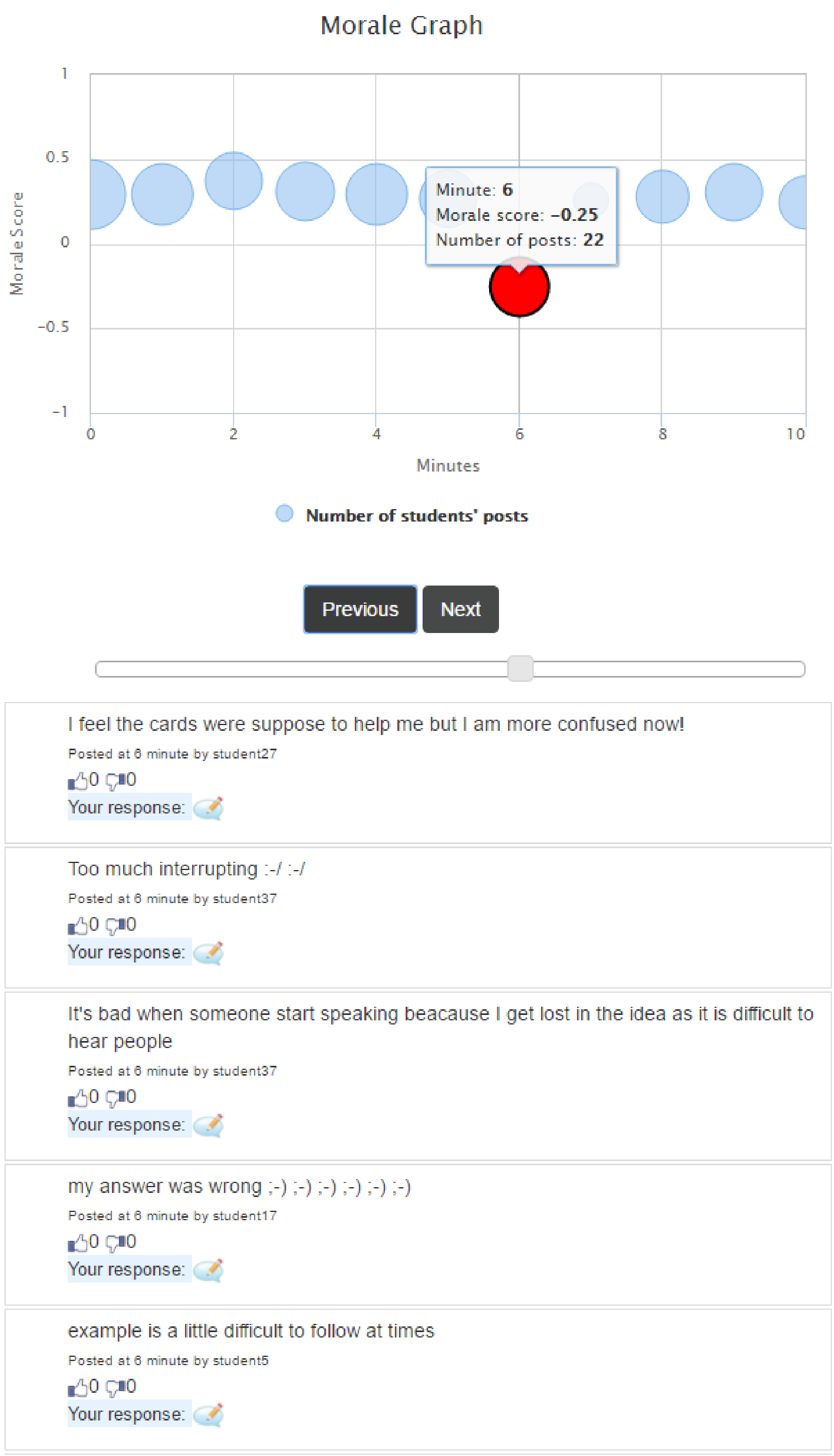

Figure 3.5: A pilot test in an Information Technology lecture

different domains. These words must be identified and reassessed to improve the accuracy of sentiment strength prediction in that particular context, for example, the teaching and learning domain for which *ClasSense* is designed. We used the lexicon adaptation method to revise words' sentiment scores and add new words to the original SentiStrength lexicon (Thelwall, 2013). There are four main steps in customising the SentiStrength lexicon using the lexicon adaptation method: (1) collecting the data, (2) labelling data by human experts, (3) checking SentiStrength classification of each text against a human code, and (4) adjusting the top 50 domain-specific terms in the lexicon.

3.6.1 Data Collection

First, we collected a corpus of 2,143 posts on 9 entry-level Information Technology topics through our backchannel system for 4 months. Participants were 35 students studying a variety of degrees in our university. To ease the data collection process, we did not conduct the experiments in real lecture environments; instead we developed an online portal for students to provide their feedback with the *ClasSense* backchannel system while watching the 9 lecture videos at their own times.

3.6.2 Human Scoring

Then, we asked three linguistic experts to perform manual sentiment analysis and code each post with a positive strength of 1 to 5 or a negative strength of -1 to -5. The percent agreements between human coders are moderate as shown in Table 3.2. We also applied Krisppendorff's inter-coder weighted alpha to determine the inter-rater reliability of our human coders because of its reliability for annotation in studies on emotion and opinion analysis involving more than 2 human coders (Antoine et al., 2014). The Krisppendorff's alpha reliability estimation for both positive and negative posts is close to the lowest conceivable limit ($\alpha \geq .667$) (Krippendorff, 2004), but it is still acceptable (Lombard et al., 2010). In addition, the Fleiss kappa for both positive and negative posts were calculated and had a fair strength of agreement,between 0.21-

0.40, as suggested in (Landis and Koch, 1977).

Table 3.2: Percent agreement, Krippendorff's alpha and Fleiss kappa between 3 human coders

	Positive	Negative
Coder 1 vs. 2	32.52%	59.59%
Coder 1 vs. 3	47.88%	60.15%
Coder 2 vs. 3	49.84%	66.12%
Mean	43.41%	61.95%
Krippendorff's alpha	0.6397	0.6475
Fleiss kappa	0.222	0.313

3.6.3 Adjustment of Terms

After receiving human-classified posts, we randomly selected half of the corpus (1,072 posts) whose human scores were used as the ground truth to compare against SentiStrength scores. Table 3.3 and Table 3.4 shows some of the positive and negative posts that received different scores from human and SentiStrength with original lexicon.

Then, we used SentiStrength *termWeights* command to record each disagreement term in the corpus and calculate the term's frequency as well as the average difference between its SentiStrength and human scores for both positive and negative sentiments. The results were used to guide the adjustment of terms in the SentiStrength lexicon, including adding new terms and modifying the scores of existing terms. As such, we selected the top 50 sentimental terms,

Table 3.3: An example of positive posts that received different scores from human and SentiStrength with original lexicon

Post	[+,-] Human score	[+,-] SentiStrength score
The slides are quite clear and easy to understand :)	[4,-1]	[1,-1]
Easy understand for database procedure ;-)	[4,-1]	[1,-1]
lecturer speaks clearly	[3,-1]	[1,-1]
The explanation of these terms are really clear	[4,-1]	[1,-1]
using cards is nice way to delver the idea	[4,-1]	[2,-1]

excluding noun, stop words and non-English words, based on their frequency. For each term, we considered adjusting its score in the existing lexicon based on the calculated average difference.

Table 3.4: An example of negative posts that received different scores from human and SentiStrength with original lexicon

Post	[+,-] Human score	[+,-] SentiStrength score
Puzzles are fun in the class but scary in the exam :-O :-O :-O :-O :-O	[3,-2]	[2,-4]
Wow that's a scary simulation!	[2,-3]	[3,-5]
some slides are too wordy I think :(	[1,-3]	[1,-2]
hard to understand material	[1,-3]	[1,-1]
it's hard to see some of the diagram given on the slide	[1,-3]	[1,-1]

If it was not in the existing lexicon, then it was added to the customised lexicon with a positive score equal to the rounded-up PosClassAvDiff value (or 1) if the term was positive. Otherwise, a negative score equal to the rounded-up NegClassAvDiff value (or -1) was assigned to the term if it was negative.

However, if a term was already in the existing lexicon, its polarity in the existing lexicon was considered. If the term's polarity was positive, a value of Positive Class Average Difference (PosClassAvDiff) of the term in the results was added to the term's existing score. Otherwise, a value of Negative Class Average Difference (NegClassAvDiff) of the term in the results was added to the term's existing score instead. The new score was used in the customised lexicon. The term adjustment algorithm is summarised in the following pseudocode and Table 3.5 lists the top 50 sentimental terms with results from SentiStrength *termWeights* command, original scores and adjusted scores, where the score of a term not in the original lexicon is marked N/A.

for all *word* in *results* **do**
    if *word* not in *existingLexicon* **then**
        $customisedLexicon.add(word)$
        if *word* is *positive* **then**
            if $PosClassAvDiff.value \leq 0$ **then**
                $customisedLexicon.modify(word, score, 1)$
            else
                $customisedLexicon.modify(word, score, roundUp(PosClassAvDiff.value))$
            end if
        else if *word* is *negative* **then**

```
          if NegClassAvDiff.value ≤ 0 then
              customisedLexicon.modify(word, score, −1)
          else
              customisedLexicon.modify(word, score, roundUp(NegClassAvDiff.value))
          end if
      end if
  else if word in existingLexicon then
      if word.polarity is positive) then
          word.score ← word.score + roundUp(PosClassAvDiff.value)
          if word.score < 1 then
              word.score = 1
          else if word.score > 5 then
              word.score = 5
          end if
      else if word.polarity is negative) then
          |word.score| ← |word.score| + roundUp(NegClassAvDiff.value)
          if |word.score| < 1 then
              word.score = −1
          else if |word.score| > 5 then
              word.score = −5
          end if
      end if
  end if
end for
```

3.6.4 Evaluation of Customised Lexicon

The other half of the corpus (1,071 posts) was used to compare the classification results made by human experts, by SentiStrength with the original lexicon and by SentiStrength with the

Table 3.5: The top 50 sentimental terms from the training set

Terms	Frequency	PosClassAvDiff	NegClassAvDiff	Original Score	Adjusted Score
good	179	1.184	0.128	2	3
clear	67	1.448	0.493	N/A	1
like	65	0.754	0.385	2	3
great	45	0.600	0.133	3	4
nice	42	1.143	0.071	2	3
understand	37	1.108	0.784	N/A	1
interesting	35	1.143	0.057	2	3
well	26	2.115	0.192	1	3
easy	25	2.200	0.000	1	3
hard	24	0.250	1.792	N/A	-2
easier	17	0.706	0.588	1	2
useful	16	0.875	0.313	2	3
important	14	0.429	0.214	1	1
love	14	0.857	0.000	3	4
helpful	13	0.231	0.077	2	2
beneficial	12	0.167	1.083	N/A	1
difficult	9	0.333	0.667	-2	-3
engaging	9	2.222	0.000	1	3
bad	7	0.429	0.714	-2	-3
cool	6	1.500	0.000	2	4
lost	6	-0.167	2.167	-1	-3
relevant	6	0.833	0.667	N/A	1
stuff	6	1.000	0.500	N/A	-1
boring	5	0.000	1.800	-2	-4
ok	5	1.000	0.400	1	2
please	5	-1.200	1.000	2	1
excellent	4	0.000	-0.250	4	4
fun	4	1.500	-0.500	2	4
haha	4	0.750	0.000	2	3
inspiration	4	0.000	0.500	3	3
problem	4	0.250	0.000	-2	-2
scary	4	0.250	-1.500	-4	-2
best	3	0.000	0.000	2	2
critical	3	0.000	0.000	-2	-2
easily	3	1.333	0.000	1	2
wordy	3	0.333	2.333	N/A	-2
against	2	0.500	1.000	-2	-3
confusing	2	0.500	0.500	-2	-3
distracted	2	0.500	0.000	-2	-2
exciting	2	-0.500	-0.500	3	2
interrupting	2	0.500	0.500	-2	-3
lack	2	0.000	1.000	-2	-3
lol	2	-1.000	0.500	2	1
perfect	2	1.500	0.500	2	4
struggle	2	0.000	2.000	-2	-4
yay	2	-0.500	0.000	2	1
confident	1	-2.000	2.000	2	1
enthusiasm	1	-1.000	1.000	3	2
improve	1	-1.000	2.000	2	1
successful	1	-2.000	1.000	3	1
support	1	-1.000	0.000	2	1

customised lexicon. Performance measurement was based on Pearson correlation, precision, recall, accuracy and F1-score.

First, a scatterplots in Figure 3.6 and Figure 3.7 for positive posts and in Figure 3.8 and Figure 3.9 for negative posts indicate linear relationships between the scores produced by human and those by SentiStrength with both the original and the customised lexicon respectively. As a result, the Pearson correlation was used to measure the closeness of SentiStrength's predicted sentiment to the actual sentiment declared by human. The results reveal a significant positive relationship between the scores produced by human and those by SentiStrength with both the original and the customised lexicon (p¡.001). Table 3.6 and Table 3.7 show a comparison of Pearson correlation between human and SentiStrength with the original and the customised lexicon for positive and negative posts respectively.

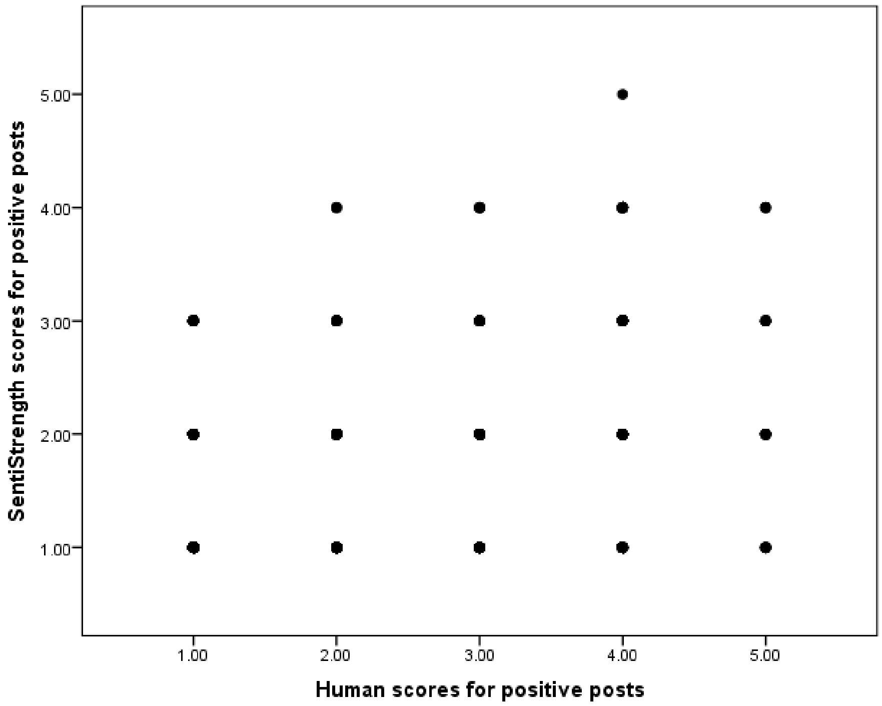

Figure 3.6: A linear relationship between human and SentiStrength scores with original lexicon for positive posts

Table 3.6: Pearson correlation for positive posts

	Human	Original Lexicon	Customised Lexicon
Human	1		
Original Lexicon	0.55	1	
Customised Lexicon	0.68	0.83	1

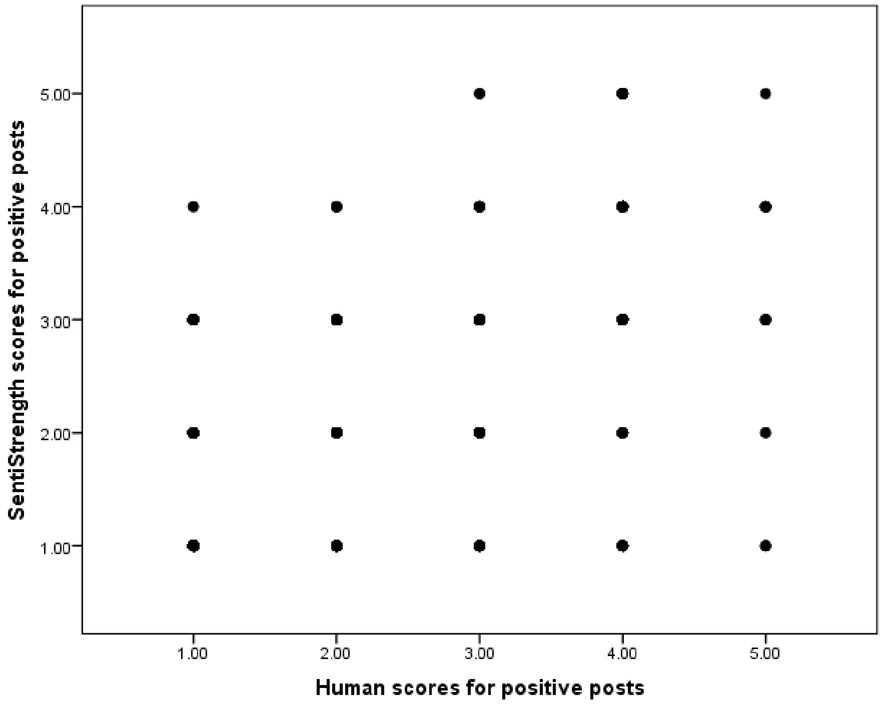

Figure 3.7: A linear relationship between human and SentiStrength scores with customised lexicon for positive posts

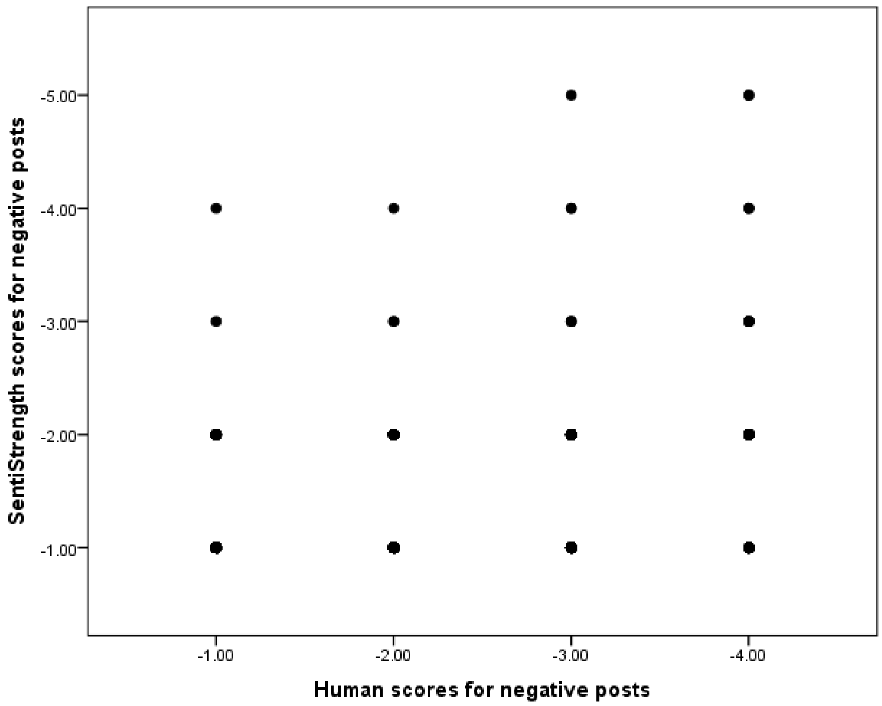

Figure 3.8: A linear relationship between human and SentiStrength scores with original lexicon for negative posts

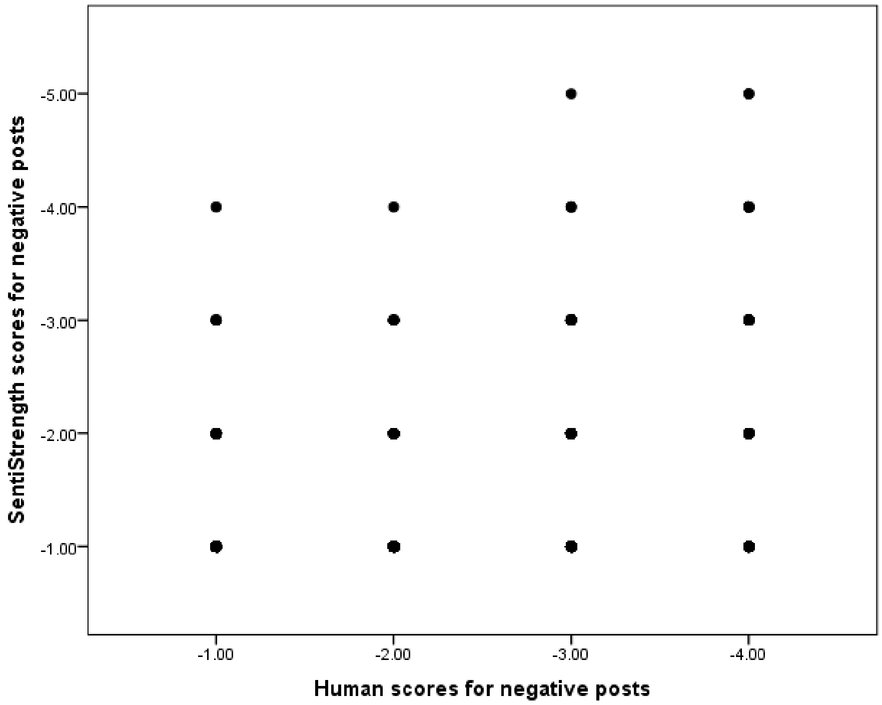

Figure 3.9: A linear relationship between human and SentiStrength scores with customised lexicon for negative posts

Table 3.7: Pearson correlation for negative posts

	Human	Original Lexicon	Customised Lexicon
Human	1		
Original Lexicon	0.40	1	
Customised Lexicon	0.49	0.83	1

The results further show that the customised lexicon achieved a 23.63% increase in correlation from 0.55 to 0.68 for positive sentiment and a 22.5% increase in correlation from 0.40 to 0.49 for negative sentiment, confirming the significance of customising the lexicon. In addition, we computed overall accuracy, precision, recall, and F1-score of SentiStrength with the original (O) and the customised (C) lexicon for both positive and negative scores by using the Scikit-learn machine learning library (Pedregosa et al., 2011). The results are summarised in a multiclass classification matrix as shown in Table 3.8, Table 3.9 and Table 3.10.

Table 3.8: Overall accuracy

	Original Lexicon		Customised Lexicon	
	Positive Score	Negative Score	Positive Score	Negative Score
Overall Accuracy	0.43	0.57	0.53	0.59

Table 3.9: Classification metrics for positive scores

Positive Scores	Precision		Recall		F1-score	
	O	C	O	C	O	C
1	0.68	**0.74**	0.83	0.81	0.75	**0.78**
2	0.13	**0.21**	0.39	0.20	0.19	**0.20**
3	0.15	**0.32**	0.09	**0.53**	0.11	**0.40**
4	0.62	**0.64**	0.06	**0.25**	0.11	**0.36**
5	0.00	**0.06**	0.00	**0.04**	0.00	**0.05**
Mean	0.48	**0.56**	0.43	**0.53**	0.39	**0.52**

Table 3.10: Classification metrics for negative scores

Negative Scores	Precision		Recall		F1-score	
	O	C	O	C	O	C
-1	0.64	**0.67**	0.96	0.95	0.77	**0.79**
-2	0.19	0.19	0.16	0.15	0.17	0.17
-3	0.25	**0.37**	0.02	**0.11**	0.03	**0.16**
-4	0.43	**0.55**	0.03	**0.06**	0.06	**0.11**
-5	0.00	0.00	0.00	0.00	0.00	0.00
Mean	0.48	**0.53**	0.57	**0.59**	0.48	**0.51**

As shown in Table 3.8, overall accuracy of SentiStrength with the customised lexicon is higher than that of SentiStrength with the original lexicon for both positive and negative sentiments. In addition, the achieved accuracy is near 60% which is the average accuracy of publicly available sentiment analysis tools (Cieliebak et al., 2014). The precision, recall and F1-score results in Table 3.9 and Table 3.10 further confirm that SentiStrength with the customised lexicon outperform SentiStrength with the original lexicon, again confirming the significance of customising the lexicon.

Chapter 4

Evaluation

Participants were 35 students studying a variety of degrees at Flinders University. We recruited the students by posting an advertisement for research participation on university billboards and through emails. The number of participating students is considered as a large class setting (Weaver and Qi, 2005). Also, 7 lecturers at School of Computer Science, Engineering and Mathematics, Flinders University were recruited by an invitation email with an approval from the Dean of School of Computer Science, Engineering and Mathematics, Flinders University.

4.1 Population

As shown in Table 4.1, the respondents made up of equally number of male and female. The majority of the respondents were between the ages of 25 or older (45.7%). There were more respondents in the 20 to 24 group (40.0%) than the 19 or younger group (14.3%). The number of international and domestic students were about the same. Most students were enrolled as full-time status (94.3%), while students enrolled part-time made up only 5.7% of the respondents. The proportion of respondents who were undergraduate and graduate students were about the same. The majority of the respondents were first year undergraduate student (n=8, 22.9%) as shown in Table 4.1.

Table 4.1: Demographic information of participants

Demographic factors		Frequency	Percentage
Gender	Male	17	48.6
	Female	18	51.4
Age	19 or younger	5	14.3
	20 to 24	14	40.0
	25 or older	16	45.7
Student type	International	17	48.6
	Domestic	18	51.4
Enrollment status	Full-time	33	94.3
	Part-time	2	5.7
Degree	Undergraduate	18	51.4
	Master	8	22.9
	PhD	9	25.7
Year of study	1	13	37.1
	2	7	20.0
	3	5	14.3
	4	7	20.0
	5	3	8.6
Course	Advanced Arts	1	2.9
	Biomedical Engineering	2	5.7
	Clinical Pharmacology	1	2.9
	Commerce	1	2.9
	Computer and Network Systems Engineering	1	2.9
	Computer Science	3	8.6
	Creative arts	1	2.9
	Disability and Developmental Education	1	2.9
	Education	1	2.9
	Electrical Engineering	1	2.9
	Electronic Engineering	1	2.9
	Engineering Technology	1	2.9
	Information Technology	4	11.4
	Mathematical Sciences	1	2.9
	Mechanical Engineering	1	2.9
	Medical Biotechnology	1	2.9
	Medicine	2	5.7
	Molecular Biology	1	2.9
	Nursing	1	2.9
	Philosophy	1	2.9
	Robotics	2	5.7
	Social Work	3	8.6
	Software Engineering	2	5.7
	Speech Pathology	1	2.9

4.2 Design and development of measurement instruments

In this research, three measurements were selected to assess satisfaction and acceptance toward using the proposed visualisation and functions of the *ClasSense* backchannel system. Firstly, the original System Usability Scale (SUS) was used as an instrument to compare lecturers satisfaction toward using the *ClasSense* backchannel system and a conventional backchannel system. Secondly, we used the Technology Acceptance Model (TAM) to predict the lecturers and students acceptance of the *ClasSense* backchannel system. Thirdly, the *Seven Principles* were used as a framework to evaluate students satisfaction toward using the *ClasSense* backchannel system.

The second and third questionnaires were developed by modifying statements from the literature related to TAM and the *Seven Principles*. The *ClasSense* backchannel system was trialled at Flinders University for 4 months, and data was collected using the following questionnaires constructs with 35 participating students and 7 lecturers at Flinders University. The following subsections introduce these measurements in more details.

4.2.1 System Usability Scale (SUS)

According to ISO 9241-11 (ISO, 1998; NCITS, 2001), usability is defined as "*Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*" To assess participants satisfaction with the perceived usability of systems after completing a set of test scenarios, standardised questionnaires have been developed with advantages such as economy, scientific generalisation, communication effectiveness (Sauro and Lewis, 2012). In addition, these measures were found to be more reliable than non-standardised usability questionnaires (Hornbæk, 2006; Hornbæk and Law, 2007; Sauro and Lewis, 2009).

There are many excellent standardised usability questionnaires such as the Questionnaire for User Interaction Satisfaction (QUIS) (Chin et al., 1988), Software Usability Measurement Inventory (SUMI) (McSweeney, 1992; Kirakowski and Corbett, 1993), Post-Study System Usability

Questionnaire (PSSUQ) (Lewis, 1992) and System Usability Scale (SUS) (Brooke et al., 1996).

Among these questionnaires, the SUS has several attributes that make it a good choice for general usability testing such as technology agnostic, quick and easy to use, providing a single score on a scale that is easy to understand for general people and non-proprietary (Bangor et al., 2008; Zviran et al., 2006). As a result, in this research, we adopted the original SUS to measure and compare usability of the *ClasSense* backchannel system that of a conventional backchannel system.

4.2.2 Technology Acceptance Model (TAM)

Another aspect that is important when introducing a new technology is to achieve user acceptance. Davis et al. (1989) proposed the Technology Acceptance Model (TAM) to predict user acceptance of Information Technology, as illustrated in Figure 4.1

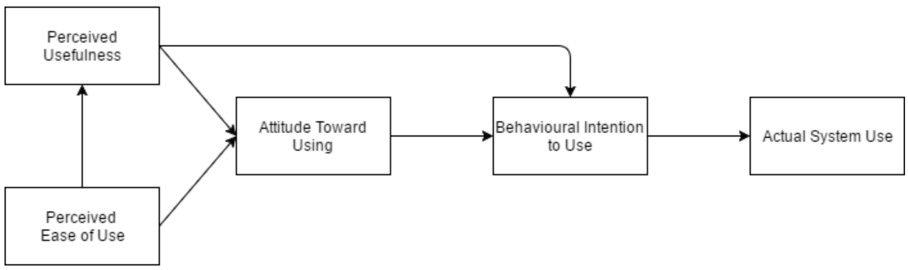

Figure 4.1: The Technology Acceptance Model (Davis et al., 1989)

The main factors of technology acceptance in TAM are perceived ease of use (PEU) and perceived usefulness (PU). For the definitions of these factors, perceived usefulness is defined as the degree to which a person believes that using the system will enhance his or her work, and perceived ease of use is the degree of ease that a person believes associated with the use of the system.

The TAM was used in numerous empirical settings (Koufaris, 2002) and validated in many works (Szajna, 1996), amongst others, to predict an acceptance of classroom response systems (e.g., (Holzer et al., 2013; Reinhardt et al., 2012; Siau et al., 2006)) and e-learning system (e.g., (Ong et al., 2004; Ngai et al., 2007; Sun et al., 2008)). In addition, the TAM is appropriate for predicting student satisfaction in using online learning systems (Arbaugh, 2000, 2002; Arbaugh and Duray, 2002; Atkinson and Kydd, 1997; Wu et al., 2006).

In addition to TAM, Venkatesh and Davis (2000) developed TAM2 in order to better understand the determinants of perceived usefulness and design organizational interventions that would increase user acceptance and usage of new systems. As illustrated in Figure 4.2, TAM2 includes additional key determinants of TAM's perceived usefulness and usage intention constructs including social influence processes (subjective norm, voluntariness, and image) and cognitive instrumental processes (job relevance, output quality, result demonstrability, and perceived ease of use).

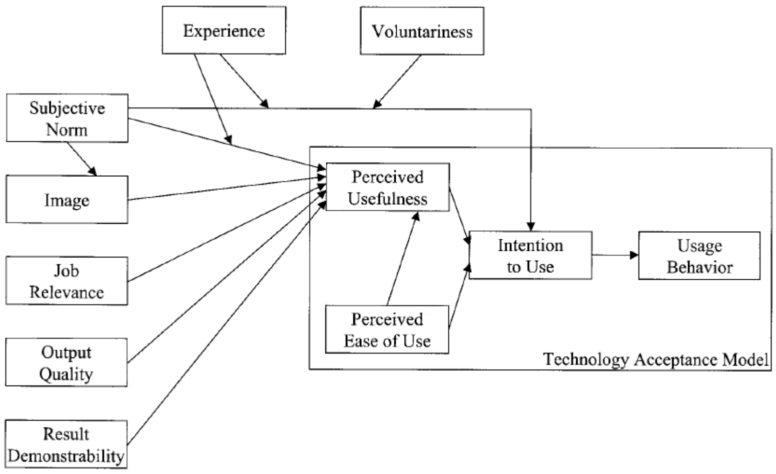

Figure 4.2: The Technology Acceptance Model 2 (Venkatesh and Davis, 2000)

Later, TAM3 was created by combining the TAM2 model with determinants of perceived ease of use including computer self-efficacy, perceptions of external control, computer anxiety, computer playfulness, perceived enjoyment and objective usability as illustrated in Figure 4.3.

In this research, we think that the acceptance and usage of *ClasSense* backchannel system depends on personal preference, not social influences. Usage of the *ClasSense* system among lecturers and students should only come from their own decision. So, the social influences presented in TAM2 are not considered in this research. In addition, we employ the standard system usability scale and a questionnaire regarding the *Seven Principle* to assess user satisfaction towards the *ClasSense* system. So, the determinants of perceived ease of use in TAM3 are not included in this research.

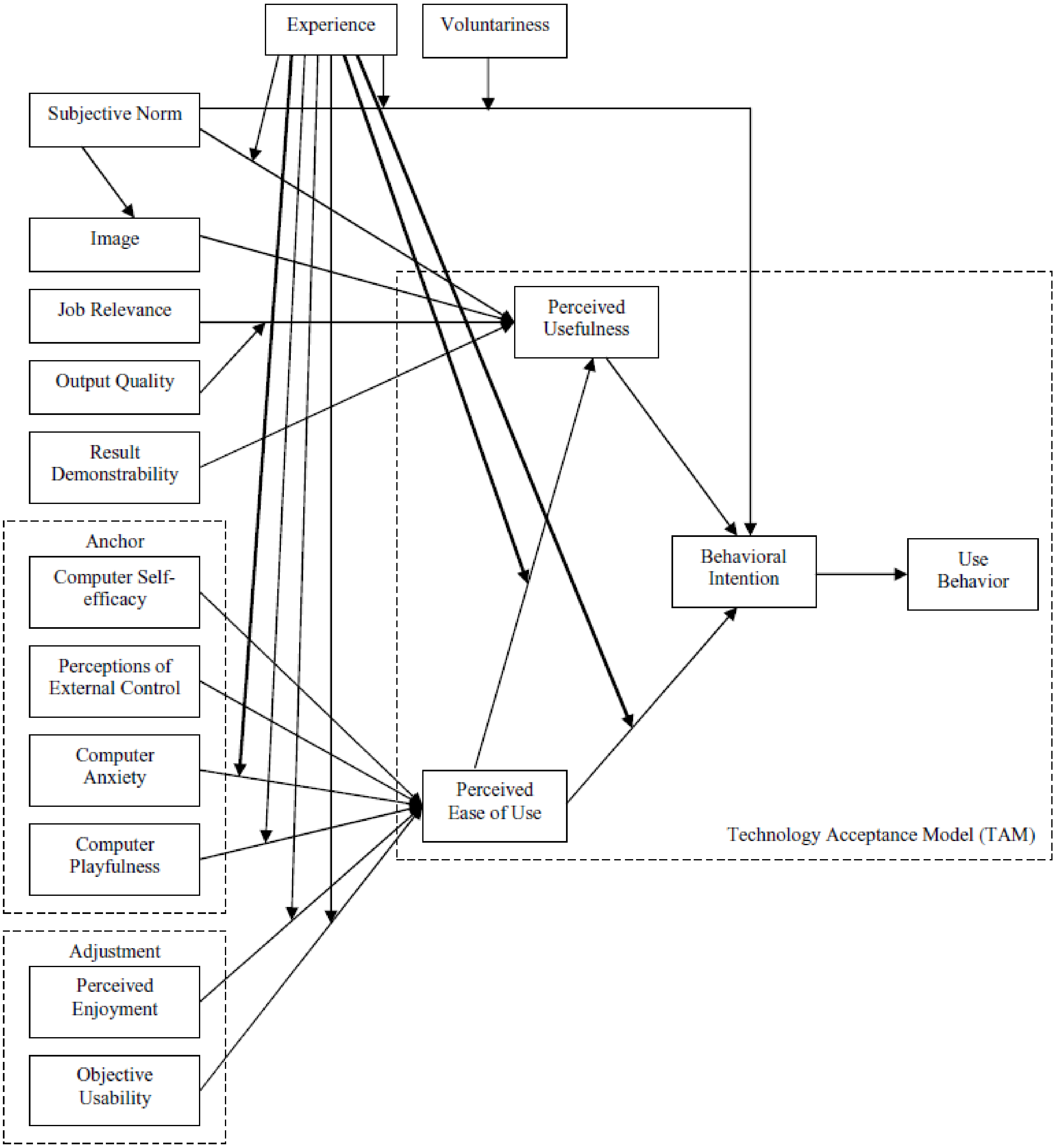

Figure 4.3: The Technology Acceptance Model 3 (Venkatesh and Bala, 2008)

As a result, we find the first version of TAM is most suitable for this research context. We used adapted items from the TAM (Davis, 1993; Davis et al., 1989; Davis, 1989; Siau et al., 2006) to predict student and lecturers acceptance of the *ClasSense* backchannel system.

4.2.3 The *Seven Principles*

In addition to TAM, we assesses students perceptions of the *ClasSense* backchannel system by using the *Seven Principles*, which is based on 50 years of research on the way lecturers teach and students learn, how students work and play with one another, and how students and lecturers talk to each other (Chickering and Gamson, 1987). The *Seven Principles* are presented in the following sections.

Encourages contact between students and faculty

Student-faculty contact has been found to be a critical factor for motivating students toward their best performance in many studies (Braxton et al., 1997; Lewis et al., 1992; Stage and Hossler, 2000). Lecturers who encourage contact with students during and after the lecture have a better opportunity to enhance student motivation, students' personal development and intellectual curiosity (Astin, 1996). In addition, lecturer characteristics such as interest in student learning, accessibility to students, friendliness, good communication skills, and enthusiasm have been recognised as having a positive impact on the relationships between students and lecturers (Marsh, 1982; Chickering and Ehrmann, 1996; Young and Shaw, 1999). Lecturers who have these attributes and provide supportive and non-threatening comments can make students feel both comfortable to approach the lecturer for help when they find difficulties in their study and confident to complete tasks and reach higher levels of achievement (Bangert, 2004).

Develops reciprocity and cooperation among students

The cooperative learning model supports the notion that social interaction promotes students' learning (Astin et al., 1993; Cooper and Mueck, 1990; Johnson et al., 1991). Students improve thinking and understanding when they have an opportunity to share and respond to each other's ideas (Chickering and Ehrmann, 1996). Working with other students has been shown to increase engagement, productivity and self-esteem (Batts et al., 2006). In addition, using computer-mediated communication (CMC) to facilitate learning is shown to provide opportunities to build cognitive and social presence beyond those afforded by face-to-face interactions (Garrison et al., 1999; Shea et al., 2003).

Uses active learning techniques

Active learning creates learning environments and experiences that bring students to discover and construct their own knowledge and solve problems (Barr and Tagg, 1995). The use of active learning has been shown to help maintain student attention and involvement (Prince, 2004), and increase students' retention of information and development of higher order learning skills, such as knowledge transfer, problem-solving measures, and motivation for further learning (McKeachie et al., 1987).

Gives prompt feedback

Effective feedback is identified as gentle, constructive and timely guidance to specific students to help them identify gaps in knowledge, goals and strategies for future learning (Getzlaf et al., 2009; Sadler, 1998; Nicol and Macfarlane-Dick, 2006). Students who receive prompt feedback know what they are doing correctly and how they can improve their work, so they can focus on their learning (Chickering and Ehrmann, 1996).

Emphasises time on task

Students' effective learning can be achieved by effective time management. (Chickering and Ehrmann, 1996). Lecturers can help students learn to use time wisely by specifying the amount of time students are expected to spend on a task when the lecturers develop course materials and learning processes (Sorcinelli, 1991).

Communicates high expectations

High expectations are important to every students (Chickering and Ehrmann, 1996). A relationship between academic achievement expectancies as predictors of the academic performance has been demonstrated in many studies (House, 1993; Tavani and Losh, 2003). Instructors who develop challenging goals, such as giving an example of good quality work that is a model of instructor expectations, can support the cognitive development that will help students evaluate future works of their newly acquired skills and knowledge (Lim and Moore, 2002).

Respects diverse talents and ways of learning

Students' growth and development in several areas such as academic, social, personal, and vocational could be facilitated by lecturers who show regard for students' unique interests and talents (Chickering and Gamson, 1991). As a result, in addition to subject matter skill, lecturers should have an understanding of the learning process and skill in facilitating individual and group learning (Sims and Sims, 1995). Various factors such as prior knowledge, personality styles, beliefs about learning, cognitive processing, and demographics of their students must be carefully considered when lecturers do instruction planning (Svinicki, 1999).

In summary, we adapted 24 items that aligned with the *Seven Principles* for Good Practice in Undergraduate Education (Chickering and Gamson, 1987) from the literature regarding students' perception on classroom response system (Laxman, 2011; Graham et al., 2007; Pelton et al., 2008; MacGeorge et al., 2008; DeBourgh, 2008; So and Brush, 2008; Porter and Tousman, 2010) and the Classroom Emotions Scale (CES) (Titsworth et al., 2010) to measure student's satisfaction with using the *ClasSense* backchannel systems in online classes.

4.3 Research questions and hypotheses

The questionnaires developed in Section 4.2 were used to collect data, answer to research questions and test hypotheses presented in this section. We used descriptive statistics to present the results and regression analysis to predict the user adoption, as explained in Section 4.5.

4.3.1 Research questions and hypotheses for student evaluation

The first research question **RQ1** was created to measure the overall student satisfaction with the *ClasSense* backchannel system in the implementation of each of the *Seven Principles*.

RQ1. What are the levels of student satisfaction toward the integration of *ClasSense* backchannel system in an online learning system?

We used descriptive statistics to answer the first research question **RQ1**. In addition, we would like to predict the students' adoption of the *ClasSense* backchannel system to use in an online learning system. The adoption was based on students' perceived usefulness, which is a perception of degrees of improvement in online learning experiences while they were using the *ClasSense* backchannel system, and perceived ease of use, which is a perception of how easy it is to use the *ClasSense* backchannel system with an online learning system. As a result, the second research question **RQ2** was created.

RQ2. How does students' perceived usefulness and ease of use affect adoption and satisfaction of using the *ClasSense* backchannel system to use in an online learning system?

To answer the second research question **RQ2**, the following hypotheses based on the diagram of the TAM model in the context of the acceptance of the *ClasSense* backchannel system were set and tested using a multiple regression analysis.

H1: Perceived usefulness and perceived ease of use will positively influence students attitude towards use of the *ClasSense* backchannel system in an online learning system.

H2: Perceived usefulness and attitude toward using will positively influence students behavioural intention to use the *ClasSense* backchannel system in an online learning system.

In addition, we hypothesised that the more students perceive usefulness and ease of use toward the *ClasSense* backchannel system features, such as giving feedback and expressing their emotions to lecturers in online learning system, the more satisfaction and chances for using the *ClasSense* backchannel system in the future. As a result, we set another hypothesis **H3** to confirm the finding from the literature (Arbaugh, 2002; Arbaugh and Duray, 2002; Pituch and Lee, 2006).

H3: Students perceived ease of use and usefulness of the *ClasSense* backchannel system will positively influence student satisfaction of the *ClasSense* backchannel system.

4.3.2 Research questions for lecturer evaluation

The fourth research question **RQ4** aims to measure how much lecturers were satisfied with the *ClasSense* backchannel system compared to a general backchannel system. We answered to this question by running paired-samples t-test to compare mean values of lecturers' responses of SUS questions for the *ClasSense* backchannel system and a general backchannel system.

RQ4. What are the levels of satisfaction toward the usability of the *ClasSense* backchannel system compared to a general backchannel system?

For **RQ5**, we would like to measure the level of lecturers' perceived usefulness and ease of use with the *ClasSense* backchannel system. We used descriptive statistics to answer to this question.

RQ5. What are the levels of perceive usefulness and ease of use for the *ClasSense* backchannel system?

4.4 Research objectives

To deal with such research questions, this study establishes three research objectives (below).

- Design and develop a web application and a backend system to support lecturer to know students' real-time morale and the current important discussions during her/his lecture

- Design and develop a backchannel system with a microblogging user interface that allows students to express their sentiments and emotions in a large lecture, and

- Design questionnaires to evaluate the usefulness, user acceptance and user satisfaction of the proposed system from the perspectives of both the lecturers and the students

4.5 Procedure

We tested the technological features, usability and user acceptance of the *ClasSense* backchannel system through an online learning portal with the participating students and lecturers at Flinders University for 4 months, during May 2016 and August 2016.

During the experiment period, participating students were asked to provide their feedback, such as asking questions and emotions expression, using the *ClasSense* backchannel system while watching the 9 lecture videos in entry-level Information Technology topics at their own times. Meanwhile, the participating topic lecturers were asked to use the morale-graph-based user interface in the *ClasSense* backchannel system to respond to the students' feedback to provide academic and affective support to student.

At the end of the experiment period, the students were asked to complete the questionnaires regarding student satisfaction and acceptance, and the lecturers were asked to use another user interface, which is commonly found in general backchannel systems, to navigate and respond to students' posts. Once the lecturers finished using each interface, they were asked to answer the same SUS questions and the lecturer acceptance questionnaire. Data from the returned surveys were recorded and processed using the Statistical Package for Social Science (SPSS).

We adopted the two steps data analysis approach proposed by Anderson and Gerbing (1988), where the first step involves the reliability testing to ensure the strength of the model, and the second step involves the correlation and regression analysis. Collected data was used in an internal consistency calculation and multiple regression analysis. The internal consistency index was informed using the Cronbach's coefficient alpha to see how reliably the questionnaire items that are designed to predict the user acceptance and satisfaction actually do so. The results of Cronbach's Alpha for the student and lecturer questionnaires are reported in Section 5.1.1 and 6.1.1 respectively. Then, the descriptive statistics were used and the results are presented in Section 5.1.3 and 6.1.2.

After that, we tested research hypotheses by using a multiple regression analysis to determine the causal effects among the variables in the technology acceptance model, which in turn answer to one of the research questions about students adoption of the *ClasSense* backchannel system. To test the hypotheses, we created combined variables in SPSS to represent mean values of items in questionnaires as explained in Section 5.1.2 and ran a single and multiple regression analysis on the combined variables. This thesis reports the results of the regression analysis for TAM factors in the student acceptance and relation between student acceptance and satisfaction in Section 5.1.4.

4.6 Statistical analysis techniques

This section presents statistical analysis techniques used in this study.

4.6.1 Reliability testing

The consistency of constructs in a measuring instrument such as a questionnaire is called reliability. It considers whether the obtained responses are a stable indication of the respondents' views of the items in a particular instrument. One form of reliability measure developed by Kuder and Richardson (1937) is the internal consistency method, and the Cronbach's coefficient alpha which we adopted is one of the most widely used internal consistency reliability indices (Cronbach, 1951). It is most commonly used with multiple Likert questions in a questionnaire that form a scale to determine if the scale is reliable.

The Cronbach's coefficient alpha ranges between negative infinity and one (Knapp, 1991). Higher values of Cronbach's alpha indicate higher internal consistency. A benchmark value of 0.7 is commonly accepted to indicate that at least some of the items measure the same construct (George, 2011), while very high value (0.95 or higher) are not necessarily desirable, as this indicates that the items may be redundant (Streiner, 2003).

This research uses Cronbach's Alpha as a method to measure the internal consistency reliability of the questions in the questionnaires because of its wide adoption and accuracy as indicated by (Davis Jr, 1986). Nunnally Jum and Bernstein Ira (1978) and Hair (2009) recommended that Cronbach's Alpha values from 0.8 or higher is a good reliability and enough for testing the reliability in basic research.

4.6.2 Regression analysis and hypotheses testing

Linear regression is used to predict the value of a variable based on the value of another variable. The variable we want to predict is called the dependent variable (or the outcome variable) and the variable we are using to predict the other variable's value is called the independent variable (or the predictor variable). Similarly, multiple regression, an extension of simple linear regression, is used to predict the value of a variable based on the value of two or more other variables (Sapsford and Jupp, 2006).

The standardised regression coefficient (β) is used to indicate the strength of a relationship between a given predictor and output in a standardised form (Field, 2009). This coefficient should not be zero and vary between -1 and +1. The stronger relation or effect there is between the factors the closer the coefficient gets to -1 or +1. However, if the coefficient is less than zero, it shows the assumed relation is in the wrong direction (Miller and Acton, 2009; Muijs, 2010; Field, 2009). Also, it is possible to examine the significance level of the hypotheses and causal effects among the variables by using the multiple regression analysis. This confidence or significance level of the hypothesis is called the p value. The standard cut-off value for the p value in this research is $p < 0.05$, which means less than 5 in 100 chance of the error in the significance level of the hypothesis.

4.6.3 Inter-rater reliability

Apart from conducting the statistical analysis with the collected data, in the process of labelling training data, we used the inter-rater reliability to determine whether two or more coders are consistent in evaluating characteristics of a message (Hallgren, 2012). Although inter-rater reliability does not insure validity, it is important for conducting a content analysis and without a proper establishment of inter-rater reliability, the interpretation cannot be valid Neuendorf (2016).

There are several different statistical tools to determine inter-rater reliability, such as percent agreement, Holsti's method, Scott's pi, Cohen's kappa, Krippendorff's alpha (Lombard et al., 2002). However, there is no consensus on a single "best" way of assessing the inter-rater reliability. It is recommended to use at least two appropriate indices based on the characteristics of data and a number of raters (Lombard et al., 2004).

In this research, we chose to report the basic percent agreement and two formal indices of reliability, the Krippendorff's alpha (Krippendorff, 2004) and Fleiss kappa (Fleiss et al., 2003). These formal indices were chosen because they are suitable to handle both multiple class (not just positive and negative) and multiple raters (Powers, 2012; Hayes and Krippendorff, 2007). Furthermore, they can easily be calculated by using online utility ReCal3 ("Reliability Calculator for 3 or more coders") and ReCal OIR ("Reliability Calculator for Ordinal, Interval, and Ratio data") (Freelon, 2013).

Chapter 5

Results of student evaluation

We conducted the user evaluation for the *ClasSense* backchannel system with 35 students who were enrolled in various faculties at Flinders University by using the student satisfaction and student acceptance questionnaires as explained in Section 4.2.

5.1 Data analysis

5.1.1 Reliability analysis

The reliability analysis was conducted to check the internal validity and consistency of the items in student satisfaction and acceptance questionnaires. The results of overall and each TAM factors reliability analysis are presented in Table 5.1 and 5.2 respectively. A high level and exceeding the threshold of 0.8 (Nunnally Jum and Bernstein Ira, 1978; Hair, 2009) of good reliability with Cronbach's Alpha values (Cronbach, 1951) were found in all questionnaires.

Table 5.1: Overall reliability of student questionnaires

Questionnaire	Cronbach's Alpha	Number of Items
Student satisfaction questionnaire	.901	24
Student acceptance questionnaire	.921	12
Total	.946	36

Table 5.2: Reliability of each TAM factors in the student acceptance questionnaire

Factor	Cronbach's Alpha	Number of Items
Perceived Usefulness (**PU**)	.848	3
Perceived Ease of Use (**PEU**)	.892	3
Attitude Toward Using (**ATU**)	.695	3
Behavioural Intention to Use (**BIU**)	.935	3

5.1.2 Correlation analysis

After conducting the reliability analysis in SPSS, we created seven new combined variables based on Mean values of items in the student satisfaction questionnaire that are associated with each of the *Seven Principles*. Similarly, four new combined variables for Mean values of four constructs of TAM were created from the student acceptance questionnaires and two new variables for Mean values of student satisfaction and student acceptance were created. Then, we conducted a correlation analysis.

The results of correlations between each principles, as shown in Table 5.3, are generally positive and significant. In addition, the results as shown in Table 5.4 indicate the positive and significant correlations between the TAM constructs: **PU**, **PEU**, **ATU** and **BIU**. This confirms the original hypothesis made in the literature of TAM. Also, the positive correlation between the Means of two questionnaires were found, as shown in 5.5.

Table 5.3: Correlations of each of the seven principles in the student satisfaction questionnaire

Principle	P1	P2	P3	P4	P5	P6	P7
P1	1						
P2	.409*	1					
P3	.605**	.452**	1				
P4	.606**	0.249	.651**	1			
P5	.379*	-0.023	.638**	0.307	1		
P6	.599**	.365*	.541**	.530**	0.244	1	
P7	.443**	.492**	.578**	.375*	.371*	.355*	1

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Table 5.4: Correlations of student acceptance questionnaire

Factor	PU	PEU	ATU	BIU
Perceived Usefulness (**PU**)	1			
Perceived Ease of Use (**PEU**)	.466**	1		
Attitude Toward Using (**ATU**)	.635**	.566**	1	
Behavioural Intention to Use (**BIU**)	.541**	.552**	.826**	1
**. Correlation is significant at the 0.01 level (2-tailed).				

Table 5.5: Correlations of student acceptance and satisfaction questionnaire

Factor	Mean of student acceptance	Mean of student satisfaction
Mean of student acceptance	1	
Mean of student satisfaction	.839**	1
**. Correlation is significant at the 0.01 level (2-tailed).		

5.1.3 Descriptive analysis

To answer research question **RQ1**: What are the levels of student satisfaction toward the integration of *ClasSense* backchannel system in an online learning system?, a descriptive analysis was conducted and results of student satisfaction with the *ClasSense* backchannel system in the implementation of each principles of the *Seven Principles* are presented in Table 5.6 - 5.12.

P1: Encourages contacts between students and faculty

Overall, students perceived that the *ClasSense* backchannel system helped maintain interactions with the lecturer throughout their online learning experiences. The majority of students felt that the *ClasSense* backchannel system helped giving feedback to lecturers (94.3%), increased their interaction with lecturers during online lessons (88.6%), helped them express their feelings to a lecturer (82.8%) and increased a level of confidence to ask questions (80%). Students noted, "I totally agree that using *ClasSense* backchannel system will increase my interaction and opinion in the online class", "Good opportunity to interact with lecturer" and "The *ClasSense* backchannel system is a very interesting and useful tool for both students and teachers. It makes it easier to ask questions even when not confident or shy to talk.", which support the results as well. However, less than half of students felt that the *ClasSense* backchannel system helped them get the emotional support from a lecturer (45.7%).

In summary, students' satisfaction with using the *ClasSense* backchannel system to interact with lecturers in an online learning system was satisfied with the second rank (Mean = 4.02, SD = 0.54) of all the *Seven Principles* as shown in Table 5.13.

Table 5.6: Frequency table of student satisfaction of P1

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P1.1: Using *ClasSense* backchannel system has increased my opportunities to interact with my lecturer during lessons (Laxman, 2011)	0	2.9	8.6	65.7	22.9	4.09	0.66
P1.2: Using *ClasSense* backchannel system helps provide feedback on my lecturer's teaching (Graham et al., 2007)	0	0	5.7	51.4	42.9	4.37	0.60
P1.3: Using *ClasSense* backchannel system gives me confidence to ask more questions (Pelton et al., 2008)	0	5.7	14.3	34.3	45.7	4.20	0.90
P1.4: Using *ClasSense* backchannel system helps me get the emotional help and support I need from my lecturer (Titsworth et al., 2010)	2.9	14.3	37.1	34.3	11.4	3.37	0.97
P1.5: Using *ClasSense* backchannel system, I can express how I feel to my lecturer (Titsworth et al., 2010)	0	0	17.1	57.1	25.7	4.09	0.66

P2: Develops reciprocity and cooperation among students

Results from the questionnaire indicate that a high percentage of students perceived that the *ClasSense* backchannel system allowed them to be aware of other students' opinions (85.7%) and attitudes (74.3%). In addition, students responded that the *ClasSense* backchannel system created more interaction with their peers (68.6%) and helped them know their performance in relation to their peers (54.2%). Students' satisfaction with using the *ClasSense* backchannel system to interact with online classmates was satisfied with the third rank (Mean = 3.92, SD = 0.58) of all the *Seven Principles* as shown in Table 5.13.

Table 5.7: Frequency table of student satisfaction of P2

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P2.1: Using *ClasSense* backchannel system promotes opportunities for peer interaction/s with my fellow students (Laxman, 2011)	0	2.9	28.6	45.7	22.9	3.89	0.80
P2.2: Using *ClasSense* backchannel system has increased my awareness of my peers' opinions (Graham et al., 2007)	0	0	14.3	54.3	31.4	4.17	.66
P2.3: Using *ClasSense* backchannel system has increased my awareness of my peers' attitudes (Graham et al., 2007)	0	0	25.7	42.9	31.4	4.06	.77
P2.4: Using *ClasSense* backchannel system helps me understand my performance in relation to my peers (Graham et al., 2007)	0	14.3	31.4	37.1	17.1	3.57	.95

P3: Uses active learning techniques

Overall, students responses suggest that the majority of students perceived that the *ClasSense* backchannel system helped them actively exchanged ideas with their online classmates (68.6%), expressed true feelings through the *ClasSense* backchannel system (65.7%), developed new knowledge from their online classmates (62.9%) and thought clearer than in a normal online learning system without the *ClasSense* backchannel system (51.4%). However, less then half of students felt that the *ClasSense* backchannel system help them understand better (45.7%) or develop new skills from their online classmates (45.7%) during online class sessions.

To conclude, students' satisfaction with using the *ClasSense* backchannel system to actively learn in online classes was satisfied with the fifth rank (Mean = 3.54, SD = 0.62) of all the *Seven Principles* as shown in Table 5.13.

Table 5.8: Frequency table of student satisfaction of P3

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P3.1: Using *ClasSense* backchannel system helps me understand better during class time (MacGeorge et al., 2008)	2.9	22.9	28.6	31.4	14.3	3.31	1.08
P3.2: Using *ClasSense* backchannel system helps me think clearer than in regular lecture sessions (DeBourgh, 2008)	2.9	22.9	22.9	40.0	11.4	3.34	1.06
P3.3: Using *ClasSense* backchannel system makes me actively exchanged my ideas with my classmates (So and Brush, 2008)	0	5.7	25.7	48.6	20.0	3.83	.82
P3.4: Using *ClasSense* backchannel system helps me develop new skills from my classmates (So and Brush, 2008)	2.9	17.1	34.3	37.1	8.6	3.31	.96
P3.5: Using *ClasSense* backchannel system helps me develop new knowledge from my classmates (So and Brush, 2008)	0	5.7	31.4	54.3	8.6	3.66	.73
P3.6: The emotions I express through *ClasSense* backchannel system can represent my true feelings (Titsworth et al., 2010)	0	2.9	31.4	48.6	17.1	3.80	.76

P4: Gives prompt feedback

Responses to the three items of principle 4 indicated that 80% of students felt more engaged while using the *ClasSense* backchannel system. 71.5% of students perceived that the *ClasSense* backchannel system helped lecturers responded to their concerns and feelings and 60% of students felt that they got immediate feedback from lecturers in an online learning system. Students' satisfaction in using the *ClasSense* backchannel system to give and receive feedback in online classes is ranked fourth (Mean = 3.90, SD = 0.78) of all the *Seven Principles* as shown in Table 5.13.

Table 5.9: Frequency table of student satisfaction of P4

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P4.1: *ClasSense* backchannel system helps me to get immediate feedback during lessons (Laxman, 2011)	2.9	11.4	25.7	28.6	31.4	3.74	1.12
P4.2: *ClasSense* backchannel system makes me feel more engaged with my learning (Laxman, 2011)	2.9	8.6	8.6	45.7	34.3	4.00	1.03
P4.3: Using *ClasSense* backchannel system helps lecturer to respond to my concerns and feelings (Titsworth et al., 2010)	0	0	28.6	42.9	28.6	4.00	.77

P5: Emphasises time on task

Results from questions pertaining to principle 5 suggest that only 60% of students surveyed agreed that the *ClasSense* backchannel system helped them move at the right pace and maintain concentration in online classes (57.1%). Students' satisfaction with the implementation of principle 5 in the *ClasSense* backchannel system is in the sixth rank (Mean = 3.50, SD = 0.82) of all the *Seven Principles* as shown in Table 5.13.

Table 5.10: Frequency table of student satisfaction of P5

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P5.1: Using *ClasSense* backchannel system helps the class move at the right pace for me (Graham et al., 2007)	2.9	2.9	34.3	51.4	8.6	3.60	.81
P5.2: Using *ClasSense* backchannel system helps me maintain concentration in lectures (Porter and Tousman, 2010)	8.6	17.1	17.1	40	17.1	3.40	1.22

P6: Communicates high expectations

Student responses indicate that only 31.4% of students felt that using *ClasSense* backchannel system helps them get a higher grade and 57.1% of them felt more confident in online classes. To be concluded, students' satisfaction with the implementation of principle 6 in the *ClasSense* backchannel system is ranked lowest (Mean = 3.47, SD = 0.70) of all the *Seven Principles* as shown in Table 5.13.

Table 5.11: Frequency table of student satisfaction of P6

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P6.1: Using *ClasSense* backchannel system helps me to get a higher grade (MacGeorge et al., 2008)	2.9	11.4	54.3	31.4	0	3.14	.73
P6.2: Using *ClasSense* backchannel system makes me feel more confident in the class (MacGeorge et al., 2008)	0	2.9	40.0	31.4	25.7	3.80	.87

P7: Respects diverse talents and ways of learning

The majority of students agreed that the *ClasSense* backchannel system made their feedback an important part of online classes (80%) and allowed them to answer anonymously (77.2%). The implementation of principle 7 in the *ClasSense* backchannel system is ranked highest (Mean = 4.17, SD = 0.67) of all the *Seven Principles* as shown in Table 5.13.

Table 5.12: Frequency table of student satisfaction of P7

Statement	Percent					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
P7.1: Using *ClasSense* backchannel system helps make my feedback an important part of class (Graham et al., 2007)	0	5.7	14.3	51.4	28.6	4.03	.82
P7.2: Using *ClasSense* backchannel system allows me to answer anonymously (Porter and Tousman, 2010)	0	0	22.9	22.9	54.3	4.31	.83

Overall, the implementations of principle 7, 1 and 2 are the top three most satisfied functions of the *ClasSense* backchannel system. This result shows that the students liked the anonymous posting feature of the *ClasSense* backchannel system most, and they felt that using the *ClasSense* backchannel system gave them more opportunities to interact with their lecturer and classmates in an online learning system.

Table 5.13: Overall student evaluations of the *Seven Principles* of Good Practice in Undergraduate Education

Principles of Good Practice in Undergraduate Education	Mean	SD	Rank Order
P7: Respects diverse talents and ways of learning	4.17	0.67	1
P1: Encourages contacts between students and faculty	4.02	0.54	2
P2: Develops reciprocity and cooperation among students	3.92	0.58	3
P4: Gives prompt feedback	3.90	0.78	4
P3: Uses active learning techniques	3.54	0.62	5
P5: Emphasises time on task	3.50	0.82	6
P6: Communicates high expectations	3.47	0.70	7

5.1.4 Hypotheses testing

In addition, we answered research question **RQ2**: How does students' perceived usefulness and ease of use affect adoption and satisfaction of using the *ClasSense* backchannel system to use in an online learning system? by testing that the research hypotheses **H1**: Perceived usefulness and perceived ease of use will positively influence students attitude towards use of the *ClasSense* backchannel system in an online learning system and that **H2**: Perceived usefulness and attitude toward using will positively influence students behavioural intention to use the *ClasSense* backchannel system in an online system in Section 4.3.1. A multiple regression analysis was conducted to test the hypotheses using responses from the student satisfaction and acceptance questionnaires. The results are presented in Table 5.14 and 5.15.

According to the results in Table 5.14, a relationship between perceived usefulness, ease of use and attitude toward using was statistically significant. Students' perceived usefulness (β = .473, p=.002) and ease of use (β = .346, p=.020) positively influence attitude toward using, so hypothesis **H1** received a full support. This means that students' perception of the *ClasSense* backchannel system usefulness for giving feedback and emotion expression in an online learning system encourages them to use the *ClasSense* backchannel system more than the perception of how easy the *ClasSense* backchannel system is.

The strong relationship was found between attitude toward using (β = .808, p=.000) and behavioural intention to use, but there was no statistical significance between perceived usefulness and behavioural intention to use (β = .028, p=.831). As a result, **H2** received only a partial support.

Table 5.14: Multiple regression results of **H1** and **H2**

Hypotheses	Dependent variables	Independent variables	Standardised coefficient (β)	R Square	Sig.
H1	Attitude Toward Using (**ATU**)	Perceived Usefulness (**PU**)	.473	.496	.002*
		Perceived Ease of Use (**PEU**)	.346		.020*
H2	Behavioural Intention to Use (**BIU**)	Attitude Toward Using (**ATU**)	.808	.683	.000*
		Perceived Usefulness (**PU**)	.028		.831

* Standardised beta coefficient is significant at the .05 level (2-tailed).

In addition, a regression analysis was performed to examine hypothesis H3: Students perceived ease of use and usefulness of the *ClasSense* backchannel system will positively influence student satisfaction of the *ClasSense* backchannel system. As shown in Table 5.15, H3 was supported with p-values less than .05. The results revealed that students' perceived ease of use (β=.375) and usefulness (β=.396) have positive relationships with students' satisfaction. Both factors can be used to explain 75.2% of the student satisfaction's variance. This confirms the finding in the literature as explained in Section 4.3.1.

Table 5.15: Multiple regression result of H3

Hypotheses	Dependent variables	Independent variables	Standardised coefficient (β)	R Square	Sig.
H3	Student Satisfaction	Perceived Ease of Use (**PEU**)	.375	.752	.002*
		Perceived Usefulness (**PU**)	.396		.002*

* Standardised beta coefficient is significant at the .05 level (2-tailed).

5.2 Discussion

The student evaluation part of this research answers two research questions **RQ1** and **RQ2**. The first research question **RQ1** was answered by the descriptive analysis explained in Section 5.1.3.

Overall student satisfaction is high with most students reporting positive feedback such as "I really liked the application.", "This web application is a good idea to get student paying more attention for their class.", "Overall this was an interesting experience and I enjoyed testing it out" and "*ClasSense* students web application was found to be a good tool.".

To answer the research question **RQ2**, hypotheses **H1** and **H2** were set. The hypothesis **H1** is confirmed with some students reporting that they want to use the application in the future as shown below.

- "Hopefully it will be used across all universities/schools"
- "it is a nice tool that can be integrated in the future"

- "it is a better idea to run this application with real teaching environment in the future"
- "*ClasSense* web application is easy to use"
- "The *ClasSense* student web is a very interesting and useful tool for both students and teachers. It makes it easier to ask questions even when not confident or shy to talk. This program would really help student to learn and therefore in turn give them confidence to do well in the topic. It is a very useful tool for student success"

Hypothesis **H2** is partially confirmed, only the **ATU** has an influence on the **BIU**; however, this is still corresponding with the TAM model. Below is the list of some comments from students that may have an impact on the future study and improvement of the *ClasSense* student application. Some concerns from students are listed below for further improvements.

- "students in the class should use this web in the right objective such as to ask questions and share opinions. Otherwise, it can interrupt other students to concentrate lessons in the class."
- "Using *ClasSense* was clunky and distraction from the lecture. This could be improved via some minor tweak to the interface."
- "I suggest that *ClasSense* is more suitable with lab class or tutorial class than lecture class. Lecture class, student should concentrate with contents from lecturer. Listen and interact with lecture in normal way should be good."
- "It might be needed to have a control in the use of it. Otherwise the students might get distracted and stop paying attention to classes."

Hypothesis **H3** is confirmed. This is to confirm the findings from the literature that student satisfaction is influenced by student perceived ease of use and perceived usefulness.

In addition, some limitations and issues are found in this experiment.

- Many posts are more about the characteristics of video content (such as light, sound and lecture style) than the content of lecture videos. This issue may come from the fact that about half of the participants are not studying in engineering or IT, so they might not have an understanding in the topic that they are watching.

- Students seem to not using the ‘like’ and ‘dislike’ buttons as much as we thought. This issue is quite surprising to us as we found that these buttons are used a lot in social media sites like Facebook. So, we may have to do more investigations on this issue.

- Participant recruiting period takes too long than we expected.

Chapter 6

Results of lecturer evaluation

We conducted the user evaluation for the *ClasSense* backchannel system with lecturers of the School of Computer Science, Engineering and Mathematics at Flinders University by using the original System Usability Scale (SUS) (Brooke et al., 1996) and adapted items from the Technology Acceptance Model (TAM) (Davis, 1989). The scale of both questionnaires ranges from 1 (Strongly disagree) to 5 (Strongly agree). The completed questionnaires were collected from 7 lecturers. Among them, 6 were male and 1 were female.

6.1 Data analysis

6.1.1 Reliability analysis

We conducted the reliability analysis to check the internal validity and consistency of the items in lecturer acceptance questionnaires. As shown in Table 6.1, our technology acceptance questionnaire has a high level of overall reliability equal to Cronbach's Alpha (Cronbach, 1951) value of 0.864, which is exceeded the threshold of 0.8 (Nunnally Jum and Bernstein Ira, 1978; Hair, 2009).

Table 6.1: Reliability of lecturer acceptance questionnaire

Factor	Cronbach's Alpha	Number of Items
Perceived Usefulness (**PU**)	.556	4
Perceived Ease of Use (**PEU**)	.203	4
Attitude Towards Using (**ATU**)	.702	3
Behavioural Intention to Use (**BIU**)	.931	3
Total	.864	14

6.1.2 Descriptive analysis

To answer the following research questions: **RQ4** - What are the levels of satisfaction toward the usability of the ClasSense backchannel system compare to a general backchannel system? and **RQ5** - What are the levels of perceive usefulness and ease of use for the ClasSense backchannel system? as explained in Section 4.3.2, a descriptive analysis and paired-samples t-test were conducted and results of lecturer satisfaction and acceptance of the *ClasSense* backchannel system are presented in this section.

Lecturer satisfaction

To evaluate the usability satisfaction of the *ClasSense* backchannel system, we calculated means from SUS scores of the *ClasSense* backchannel system and compared with the General backchannel system. A comparison of the results from all participating lecturers are shown in Figure 6.1 and Table 6.2. A summary of the SUS statements, frequency tables of data collected from the *ClasSense* and General backchannel system SUS questionnaires are shown in Table 6.3 and Table 6.4.

In summary, 6 of 7 participating lecturers preferred using the morale-graph-based user interface in the *ClasSense* backchannel system over the traditional backchannel system with the high mean of SUS scores at 86.07, which indicates excellent usability (Mean of SUS score >80 of 100 (Bangor et al., 2008)), for the *ClasSense* backchannel system, as shown in Table 6.2.

In addition, a paired-samples t-test was run to test for statistical significance. The result show that the *ClasSense* backchannel system (M = 86.07, SD = 6.10) has significantly higher SUS scores than the General backchannel system (M = 72.86, SD = 9.94), $t(6) = -3.011$, $p = 0.024$.

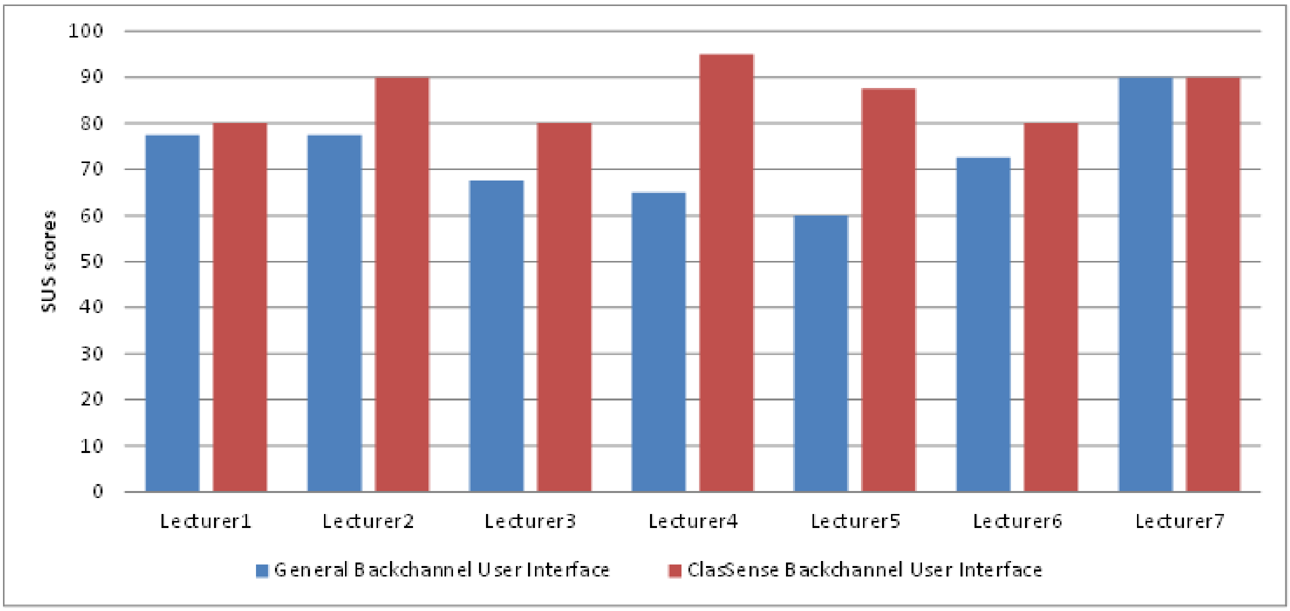

Figure 6.1: A comparison of SUS scores between the General and *ClasSense* backchannel system

Table 6.2: SUS scores between two interfaces from participating lecturers

	General Backchannel User Interface	*ClasSense* Backchannel User Interface
Lecturer1	77.50	80
Lecturer2	77.50	90
Lecturer3	67.50	80
Lecturer4	65	95
Lecturer5	60	87.50
Lecturer6	72.50	80
Lecturer7	90	90
Mean	72.86	86.07
Std. Deviation	9.94	6.10

Lecturer acceptance

In our assessment, the perceived usefulness means that using the *ClasSense* backchannel system would enable a lecturer to effectively manage a large amount of students' feedback and know class morale from an online class. Table 6.5 shows the lecturers' perceived usefulness of the *ClasSense* backchannel system that can be further categorised into four aspects. These aspects are how useful the *ClasSense* backchannel system is to a lecturer in terms of reading a large amount of students' feedback, how useful the *ClasSense* backchannel system is to a lecturer in terms of viewing the class morale graph, how useful the *ClasSense* backchannel system is

Table 6.3: Frequency table of the data collected from the *ClasSense* backchannel system SUS questionnaires used in this study

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
Q1: I think that I would like to use this system frequently.	0	0	0	6	1	4.14	0.38
Q2: I found the system unnecessarily complex.	3	3	1	0	0	1.71	0.76
Q3: I thought the system was easy to use.	0	0	0	3	4	4.57	0.53
Q4: I think that I would need the support of a technical person to be able to use this system.	6	1	0	0	0	1.14	0.38
Q5: I found the various functions in this system were well integrated.	0	0	3	4	0	3.57	0.53
Q6: I thought there was too much inconsistency in this system.	4	2	1	0	0	1.57	0.79
Q7: I would imagine that most people would learn to use this system very quickly.	0	0	0	4	3	4.43	0.53
Q8: I found the system very cumbersome to use.	4	2	1	0	0	1.57	0.79
Q9: I felt very confident using the system.	0	0	0	1	6	4.86	0.38
Q10: I needed to learn a lot of things before I could get going with this system.	6	1	0	0	0	1.14	0.38
Average	2.3	0.9	0.6	1.8	1.4	2.87	0.54

Table 6.4: Frequency table of the data collected from the General backchannel system SUS questionnaires used in this study

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
Q1: I think that I would like to use this system frequently.	0	1	5	1	0	3.00	0.58
Q2: I found the system unnecessarily complex.	3	2	1	1	0	2.00	1.15
Q3: I thought the system was easy to use.	0	0	1	4	2	4.14	0.69
Q4: I think that I would need the support of a technical person to be able to use this system.	7	0	0	0	0	1.00	0.00
Q5: I found the various functions in this system were well integrated.	0	0	6	1	0	3.14	0.38
Q6: I thought there was too much inconsistency in this system.	4	1	1	1	0	1.86	1.21
Q7: I would imagine that most people would learn to use this system very quickly.	0	0	1	2	4	4.43	0.79
Q8: I found the system very cumbersome to use.	1	0	2	4	0	3.29	1.11
Q9: I felt very confident using the system.	0	1	0	2	4	4.29	1.11
Q10: I needed to learn a lot of things before I could get going with this system.	4	2	0	1	0	1.71	1.11
Average	1.9	0.7	1.7	1.7	1	2.89	0.81

to a lecturer in terms of viewing the most popular feedback, and how useful the *ClasSense* backchannel system is to a lecturer in terms of reviewing the history of feedback.

The high average mean value (Mean = 3.89, SD = 0.90) of perceived usefulness of the *ClasSense* backchannel system suggests that the lecturers were highly satisfied with the proposed functionalities of the *ClasSense* backchannel system and felt that the system was useful.

Table 6.5: Lecturer Perceived Usefulness

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
ClasSense lecturer web application allows me to read students' feedback effectively.	0	0	0	5	2	4.29	0.49
ClasSense lecturer web application allows me to view the class morale.	0	0	1	4	2	4.14	0.69
ClasSense lecturer web application allows me to view the most popular feedback quickly.	1	2	2	0	2	3.00	1.53
Reviewing the history of feedback in *ClasSense* lecturer web application is easy.	0	0	2	2	3	4.14	0.90
Average	0.25	0.5	1.2	2.75	2.25	3.89	0.90

In addition, the other predictor, called perceived ease of use, means how lecturers perceive that it would be easy for them to become skilful at using the *ClasSense* backchannel system. The perceived ease of use of the *ClasSense* backchannel system can be further categorised into four aspects. These aspects include how easily a lecturer can learn to operate the *ClasSense* backchannel system, how easily a lecturer can get the *ClasSense* backchannel system to do what he or she wants it to do, how easily it is to be good at using the *ClasSense* backchannel system and how easily it is to use the *ClasSense* backchannel system in overall.

As shown in Table 6.6, the high average mean value (Mean = 4.23, SD = 0.74) of the four perceived ease of use aspects of the *ClasSense* backchannel system indicates that the lecturers felt that the system was easy to use.

According to Table 6.7, attitude toward using also has a high average mean value (Mean = 3.95, SD = 0.83) to confirm that the lecturers had positive feelings about using the *ClasSense* backchannel system.

However, as shown in Table 6.8, the average mean value of lecturers' behavioural intention to use (Mean = 3.62, SD = 0.72) indicates that the lecturers was willing to use the *ClasSense*

Table 6.6: Lecturer Perceived Ease of Use

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
Learning to operate *ClasSense* lecturer web application is easy for me.	0	0	0	4	3	4.43	0.54
I find it easy to get *ClasSense* lecturer web application to do what I want it to do.	0	1	2	2	2	3.71	1.11
It is easy for me to become skilful at using *ClasSense* lecturer web application.	0	0	1	3	3	4.29	0.76
I find *ClasSense* lecturer web application easy to use.	0	0	0	4	3	4.50	0.54
Average	0	0.25	0.75	3.25	2.75	4.23	0.74

Table 6.7: Lecturer Attitude Toward Using

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
Using *ClasSense* lecturer web application is a good idea.	0	0	0	4	3	4.43	0.54
ClasSense lecturer web application makes work more interesting.	1	0	1	4	1	3.57	1.27
I like working with *ClasSense* lecturer web application.	0	0	2	4	1	3.86	0.69
Average	0.33	0	1	4.33	2.33	3.95	0.83

backchannel system at a certain level that was above the average level.

Table 6.8: Lecturer Behavioural Intention to Use

Statement	Frequency					Mean	SD
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree		
I intend to use *ClasSense* lecturer web application in the future.	0	0	5	1	1	3.43	0.79
I expect that I would use *ClasSense* lecturer web application in the future.	0	0	1	5	1	4.00	0.58
I plan to use *ClasSense* lecturer web application in the future.	0	0	5	1	1	3.43	0.79
Average	0	0	3.67	2.33	1	3.62	0.72

As described above, the average mean value of perceived usefulness, the perceived ease of use, the attitude toward using and the behavioural intention to use are between 3.62 (out of 5) and 4.23 (out of 5), which suggests that the lecturers believed that using the *ClasSense* backchannel system was useful to their work and easy to use. They also had relatively positive feelings about using the *ClasSense* backchannel system and above the average level for willingness to use the *ClasSense* backchannel system.

6.2 Discussion

The evaluation of *ClasSense* lecturer application set out to investigate two research questions **RQ4** and **RQ5**. Due to a very small number of participants, we report only the descriptive statistics as shown in Section 6.1.2 and 6.1.2.

Overall, the lecturers were more satisfied with the *ClasSense* system than the general user interface of backchannel system with some good feedback as shown below.

- "I would absolutely be thrilled to have such a tool at my disposal!"
- "The *ClasSense* interface was much better. It is graphically presented and contained better information as well."
- "It would be interesting to see it live and how that would work."

For future improvement of the application, some of them requested modification to the user interface as shown below.

- "Users should be able to sort posts on different criteria (e.g.,lowest morale first and number of likes)"
- "The application should display individual morale scores or variance and their responses on each post"
- "The application should have buttons for going back to the top of page, buttons for going to the previous and next minute, and ability to 'like' a post."

Interestingly, one of the lecturers mentioned that the future is online learning using lecture videos because class participation is low, and she thought that both morale graph and comments are important, but she would be distracted by looking at the graph during a lecture. So, she would prefer to look at the morale graph after class to modify her teaching for future lecture. In addition, she thought that students should be able to see the graph during a lecture as well.

It would be interesting to know whether the students will be more engaged or not if they can see the graph and see how many people think the same way as or different from them. However, she thought that students don't like taking note, so she would like to know how many students will use the *ClasSense* application if the university provides this system as a facility in a real class.

This experiment is limited by the number of participating lecturers and experimental environment. We first approached 17 lecturers at the beginning of the experiment and 9 of them have signed a consent form, but only 7 of them have stayed to the completion of the study. We found it difficult to attract the lecturers to join the study because of many reasons such as high workload, their topic has changed or disagreement with this kind of experiment. For example, one of the lecturers mentioned that this experiment is not fair for him because the participating students were not his students, so he didn't see any reasons to answer their feedback. In addition, we have to use the lecture videos because we don't want to interfere while the lectures are progressing and because all live lectures are video recorded by the university.

Chapter 7

Future Work

It is foreseeable that this research will be advancing in two separate areas: usability and data analytics. This chapter introduces some ideas that could be developed to improve the *ClasSense* system both student and lecturer applications.

7.1 Lecturer application

For the *ClasSense* lecturer application, there are two interesting areas to further development. The features that could be integrated into the application for enhancing the usability and performance of the application. However, it depends on the type of lecture and situations that the application will be used.

Firstly, improving on the visualisation of sentiment and emotion is probably one of the obvious features to be mentioned. It could be modified by adding more widgets to display students' emotions and sentiments in more details. For example, using a circular gauge to display a live level of sentiment (positive and negative) and using a radar graph to show a distribution of emotions. However, the application should allow lecturers to choose what they want or don't want to see in different context. For example, during a face-to-face lecture, they might want to see only a small number of widget (or no widget at all, but the system is still running to collect students' feedback for after class analysis) because they want to pay more attention to teaching and observing students' behaviours or expressions.

Secondly, a continuous performance improvement of sentiment analysis in the *ClasSense* system could be carried out by using more input data to train the sentiment analysis module and a combination of machine learning and lexicon-based approach could help improve the performance as well. This could be done as a long term process in a background as long as the application is up and running.

7.2 Student application

Usability of the *ClasSense* student application could be improved by making use of speech-to-text analysis module and augmented reality (AR) techniques.

Normally, students often have a chat or whisper with their close friends or group members, who often sit in their vicinity during a lecture. We could utilise this kind of chat as a feedback to a lecturer. Students can whisper to an app on their smartphone in order to ask a question or give feedback on the lecture. Then, the app forwards the student's whisper to the speech-to-text analysis system on a server to convert to texts and save for further analyse during or after a lecture. By using this approach, the students don't need to type on their phone or laptop while listening to the lecture. So, they can keep their concentration on the lecture and only use the app when they want to ask a question or give some feedback.

Augmented reality devices, such as Google Glass and Microsoft HoloLens, could be used in a classroom in many ways that facilitate students' learning experiences. For example, it could be used to display emotions, sentiments or interesting questions of students in a classroom, so students could know what others are thinking and feeling or what additional questions they should ask without looking at other devices. In short, there are many possibilities to explore in this interesting area in the domain of learning and teaching in traditional and online classrooms.

Appendix A

Student Questionnaires

The questionnaire regarding student satisfaction and acceptance in using the ClasSense student web application

Section A: Research information

The objective

The questionnaire is to survey student satisfaction and acceptance in using ClasSense student web application. You will be asked about your perception on the level of satisfaction and acceptance in using the ClasSense student web application. Data from this survey will be analysed to determine improvement of the ClasSense student web application.

Questionnaire instruction

We would be grateful if you could spend 15 – 20 minutes of your time to complete the following short questionnaire about your satisfaction and acceptance in using the ClasSense student web application. There is no right or wrong answer for each question. However, it is important for you to respond as accurately as possible by checking the most appropriate response. Your responses are confidential and only anonymous comments and aggregate results will be disclosed.

Section B: Personal details

Gender: ☐ Male ☐ Female

Age group: ☐ 19 years and younger ☐ 20-24 years ☐ 25 years and older

Student type: ☐ International ☐ Domestic

Enrolment status: ☐ Full-time ☐ Part-time

Year: ☐ 1 ☐ 2 ☐ 3 ☐ 4 Other:____

Course of study:

☐ Agricultural Engineering
☐ Biomedical Engineering
☐ Civil Engineering
☐ Computer Science
☐ Computer and Network Systems Engineering
☐ Design and Innovation
☐ Digital Health Systems
☐ Digital Media
☐ Electrical Engineering
☐ Electronic Engineering
☐ Information Technology
☐ Mathematical Sciences
☐ Mechanical Engineering
☐ Network and Cybersecurity Systems
☐ Naval Architecture
☐ Robotics
☐ Simulation and Serious Games
☐ Software Engineering
☐ Engineering with Science Combined Degrees
☐ Engineering Technology
☐ Engineering Science
☐ Mathematics and Science
☐ Other:___________________________

Section C: Student satisfaction of the ClasSense student web application

Please rate your level of agreement with the following statements about the ClasSense student web application use by circling one number from 1 to 5.

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Using ClasSense student web application has increased my opportunities to interact with my lecturer during lessons.	1	2	3	4	5
2. Using ClasSense student web application helps provide feedback on my lecturer's teaching.	1	2	3	4	5
3. Using ClasSense student web application gives me confidence to ask more questions.	1	2	3	4	5
4. Using ClasSense student web application helps me get the emotional help and support I need from my lecturer.	1	2	3	4	5
5. Using ClasSense student web application, I can express how I feel to my lecturer.	1	2	3	4	5
6. Using ClasSense student web application promotes opportunities for peer interaction/s with my fellow students.	1	2	3	4	5
7. Using ClasSense student web application has increased my awareness of my peers' opinions.	1	2	3	4	5
8. Using ClasSense student web application has increased my awareness of my peers' attitudes.					
9. Using ClasSense student web application helps me understand my performance in relation to my peers.	1	2	3	4	5
10. Using ClasSense student web application helps me understand better during class time.	1	2	3	4	5
11. Using ClasSense student web application helps me think clearer than in regular lecture sessions.	1	2	3	4	5

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
12. Using ClasSense student web application makes me actively exchanged my ideas with my classmates.	1	2	3	4	5
13. Using ClasSense student web application helps me develop new skills from my classmates.	1	2	3	4	5
14. Using ClasSense student web application helps me develop new knowledge from my classmates.					
15. The emotions I express through ClasSense student web application can represent my true feelings.	1	2	3	4	5
16. ClasSense student web application helps me to get immediate feedback during lessons.	1	2	3	4	5
17. ClasSense student web application makes me feel more engaged with my learning.	1	2	3	4	5
18. Using ClasSense student web application helps lecturer to respond to my concerns and feelings.	1	2	3	4	5
19. Using ClasSense student web application helps the class move at the right pace for me.	1	2	3	4	5
20. Using ClasSense student web application helps me maintain concentration in lectures.	1	2	3	4	5
21. Using ClasSense student web application helps me to get a higher grade.	1	2	3	4	5
22. Using ClasSense student web application makes me feel more confident in the class.	1	2	3	4	5
23. Using ClasSense student web application helps make my feedback an important part of class.	1	2	3	4	5
24. Using ClasSense student web application allows me to answer anonymously.	1	2	3	4	5

Section D: Student acceptance of the ClasSense student web application

Please rate your level of agreement with the following statements about the level of acceptance of ClasSense student web application by circling one number from 1 to 5.

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. It is easy for me to become skilful at using ClasSense student web application.	1	2	3	4	5
2. I find it easy to get ClasSense student web application to work as intended.	1	2	3	4	5
3. I find ClasSense student web application easy to use.	1	2	3	4	5
4. Using ClasSense student web application increases my interaction in the class.	1	2	3	4	5
5. Using ClasSense student web application makes it easier for me to interact in the class.	1	2	3	4	5
6. I find ClasSense student web application useful in enhancing my interaction in the class.	1	2	3	4	5
7. Using ClasSense student web application is a good idea.	1	2	3	4	5
8. ClasSense student web application makes lecture more interesting.	1	2	3	4	5
9. I like lectures with ClasSense student web application.	1	2	3	4	5
10. I intend to use the ClasSense student web application in the future.	1	2	3	4	5
11. I expect that I will use ClasSense student web application in the future.	1	2	3	4	5
12. I plan to use ClasSense student web application in the future.	1	2	3	4	5

Please write any comments you would like to make in the following space.

Thank you very much for your cooperation in completing this questionnaire.

Appendix B

Lecturer Questionnaires

The questionnaire regarding lecturer acceptance of the ClasSense lecturer web application

Section A: Research information

The objective

This questionnaire surveys lecturers' acceptance of ClasSense lecturer web application. You will be asked about your perception of the level of perceived usefulness, perceived ease of use, attitude toward using and behavioural intention to use of ClasSense lecturer web application. Data from this survey will be analysed to determine user acceptance of ClasSense lecturer web application.

Questionnaire instruction

We would be grateful if you could spend 15 – 20 minutes of your time to complete the following short questionnaire about your acceptance of the ClasSense lecturer web application. There is no right or wrong answer for each question. However, it is important for you to respond as accurately as possible by checking the most appropriate response. Your responses are confidential and only anonymous comments and aggregate results will be disclosed.

Section B: Lecturer acceptance of the ClasSense lecturer web application

Please rate your level of agreement with the following statements about each feature of the ClasSense lecturer web application by circling one number from 1 to 5.

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. ClasSense lecturer web application allows me to read students' feedback effectively.	1	2	3	4	5
2. ClasSense lecturer web application allows me to view the class morale.	1	2	3	4	5
3. ClasSense lecturer web application allows me to view the most popular feedback quickly.	1	2	3	4	5
4. Reviewing the history of feedback in ClasSense lecturer web application is easy.	1	2	3	4	5
5. Learning to operate ClasSense lecturer web application is easy for me.	1	2	3	4	5
6. I find it easy to get ClasSense lecturer web application to do what I want it to do.	1	2	3	4	5
7. It is easy for me to become skilful at using ClasSense lecturer web application.	1	2	3	4	5
8. I find ClasSense lecturer web application easy to use.	1	2	3	4	5
9. Using ClasSense lecturer web application is a good idea.	1	2	3	4	5
10. ClasSense lecturer web application makes work more interesting.	1	2	3	4	5
11. I like working with ClasSense lecturer web application.	1	2	3	4	5
12. I intend to use ClasSense lecturer web application in the future.	1	2	3	4	5
13. I expect that I would use ClasSense lecturer web application in the future.	1	2	3	4	5
14. I plan to use ClasSense lecturer web application in the future.	1	2	3	4	5

Please write any comments you would like to make in the following space.

Thank you very much for your cooperation in completing this questionnaire.

Appendix C

System Usability Scale Questionnaires

The Questionnaire regarding the System Usability Scale of the ClasSense Lecturer Web Application

Please rate your level of agreement with the following statements about the ClasSense Lecturer Web Application by circling one number from 1 to 5.

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think that I would like to use this system frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

The Questionnaire regarding System Usability Scale of the General Backchannel System

Please rate your level of agreement with the following statements about the General Backchannel System by circling one number from 1 to 5.

Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think that I would like to use this system frequently.	1	2	3	4	5
2. I found the system unnecessarily complex.	1	2	3	4	5
3. I thought the system was easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system.	1	2	3	4	5
5. I found the various functions in this system were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this system.	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
8. I found the system very cumbersome to use.	1	2	3	4	5
9. I felt very confident using the system.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

Bibliography

H. Aagard, K. Bowen, and L. Olesova. Hotseat: Opening the backchannel in large lectures. *Educause Quarterly*, 33(3), 2010.

M. Abdous and C.-J. Yen. A predictive study of learner satisfaction and outcomes in face-to-face, satellite broadcast, and live video-streaming learning environments. *The Internet and Higher Education*, 13(4):248–257, 2010.

H. Abelson. The creation of opencourseware at mit. *Journal of Science Education and Technology*, 17(2):164–174, 2008.

A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau. Sentiment analysis of twitter data. In *Proceedings of the workshop on languages in social media*, pages 30–38. Association for Computational Linguistics, 2011.

W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of time-oriented data.* Springer Science & Business Media, 2011.

E. Airoldi, X. Bai, and R. Padman. Markov blankets and meta-heuristics search: sentiment extraction from unstructured texts. In *International Workshop on Knowledge Discovery on the Web*, pages 167–187. Springer, 2004.

F. Aisopos, G. Papadakis, and T. Varvarigou. Sentiment analysis of social media content using n-gram graphs. In *Proceedings of the 3rd ACM SIGMM international workshop on Social media*, pages 9–14. ACM, 2011.

I. E. Allen and J. Seaman. *Learning on demand: Online education in the United States, 2009.* ERIC, 2010.

M. Ally. Using learning theories to design instruction for mobile learning devices. *Mobile learning anytime everywhere*, pages 5–8, 2005.

N. Altrabsheh, M. Gaber, and M. Cocea. Sa-e: sentiment analysis for education. In *5th KES International Conference on Intelligent Decision Technologies*, 2013.

S. Amatya and A. Kurti. Cross-platform mobile development: challenges and opportunities. In *ICT Innovations 2013*, pages 219–229. Springer, 2014.

J. C. Anderson and D. W. Gerbing. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3):411, 1988.

R. J. Anderson, R. Anderson, T. VanDeGrift, S. Wolfman, and K. Yasuhara. Promoting interaction in large classes with computer-mediated feedback. In *Designing for change in networked learning environments*, pages 119–123. Springer, 2003.

G. Anthes. Html5 leads a web revolution. *Communications of the ACM*, 55(7):16–17, 2012.

J.-Y. Antoine, J. Villaneau, and A. Lefeuvre. Weighted krippendorff's alpha is a more reliable metrics for multi- coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation. In *EACL 2014*, page 10 p., 2014.

J. B. Arbaugh. Virtual classroom characteristics and student satisfaction with internet-based mba courses. *Journal of management education*, 24(1):32–54, 2000.

J. B. Arbaugh. Managing the on-line classroom: A study of technological and behavioral characteristics of web-based mba courses. *The Journal of High Technology Management Research*, 13(2):203–223, 2002.

J. B. Arbaugh and R. Duray. Technological and structural characteristics, student learning and satisfaction with web-based courses an exploratory study of two on-line mba programs. *Management learning*, 33(3):331–347, 2002.

M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al. A view of cloud computing. *Communications of the ACM*, 53(4): 50–58, 2010.

A. R. Artino. Online or face-to-face learning? exploring the personal factors that predict students' choice of instructional format. *The Internet and Higher Education*, 13(4):272–276, 2010.

A. R. Artino. Emotions in online learning environments: Introduction to the special issue. *The Internet and Higher Education*, 15(3):137–140, 2012.

A. W. Astin. Involvement in learning revisited: Lessons we have learned. *Journal of college student development*, 37(2):123–34, 1996.

A. W. Astin et al. *What matters in college?: Four critical years revisited*, volume 1. Jossey-Bass San Francisco, 1993.

M. Atkinson and C. Kydd. Individual characteristics associated with world wide web use: an empirical study of playfulness and motivation. *ACM SIGMIS Database*, 28(2):53–62, 1997.

S. Baccianella, A. Esuli, and F. Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204, 2010.

X. Bai. Affordance of ubiquitous learning through cloud computing. In *2010 Fifth International Conference on Frontier of Computer Science and Technology*, pages 78–82. IEEE, 2010.

A. W. Bangert. The seven principles of good practice: A framework for evaluating on-line teaching. *The Internet and Higher Education*, 7(3):217–232, 2004. ISSN 1096-7516.

A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

H. Bär, E. Tews, and G. Rößling. Improving feedback and classroom interaction using mobile phones. *Proceedings of Mobile Learning*, pages 55–62, 2005.

L. Barbosa and J. Feng. Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 36–44. Association for Computational Linguistics, 2010.

R. B. Barr and J. Tagg. From teaching to learninga new paradigm for undergraduate education. *Change: The magazine of higher learning*, 27(6):12–26, 1995.

M. S. Bartlett, G. C. Littlewort, M. G. Frank, C. Lainscsek, I. R. Fasel, and J. R. Movellan. Automatic recognition of facial actions in spontaneous expressions. *Journal of multimedia*, 1(6):22–35, 2006.

D. Batts, S. M. Colaric, and C. McFadden. Online courses demonstrate use of seven principles. *International Journal of Instructional Technology and Distance Learning*, 3(12):15–25, 2006.

W. Beekes. The millionairemethod for encouraging participation. *Active learning in higher education*, 7(1):25–36, 2006.

T. Bergstrom, A. Harris, and K. Karahalios. Encouraging initiative in the classroom with anonymous feedback. In *IFIP Conference on Human-Computer Interaction*, pages 627–642. Springer, 2011.

A. Bermingham and A. F. Smeaton. Classifying sentiment in microblogs: is brevity an advantage? In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1833–1836. ACM, 2010.

C. Biever. Twitter mood maps reveal emotional states of america. *New Scientist*, 207(2771): 14, 2010.

H. Bin. The study of mobile education development based on 3g technique and cloud computing. In *Uncertainty Reasoning and Knowledge Engineering (URKE), 2011 International Conference on*, volume 1, pages 86–89. IEEE, 2011.

H. H. Binali, C. Wu, and V. Potdar. A new significant area: Emotion detection in e-learning using opinion mining techniques. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies*, pages 259–264. IEEE, 2009.

M. M. Bradley and P. J. Lang. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Technical report C-1, the center for research in psychophysiology, University of Florida, 1999.

J. M. Braxton, A. V. Shaw Sullivan, and R. M. Johnson. Appraising tinto's theory of college student departure. *HIGHER EDUCATION-NEW YORK-AGATHON PRESS INCORPORATED-*, 12:107–164, 1997.

J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189 (194):4–7, 1996.

D. Carless. Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2): 219–233, 2006. ISSN 0307-5079.

S. Cetintas, L. Si, H. P. Aagard, K. Bowen, and M. Cordova-Sanchez. Microblogging in a classroom: Classifying students' relevant and irrelevant questions in a microblogging-supported classroom. *IEEE Transactions on Learning Technologies*, 4(4):292–300, 2011.

W. Chamlertwat, P. Bhattarakosol, T. Rungkasiri, and C. Haruechaiyasak. Discovering consumer insight from twitter via sentiment analysis. *J. UCS*, 18(8):973–992, 2012.

C.-M. Chen and T.-H. Lee. Emotion recognition and communication for reducing second-language speaking anxiety in a web-based one-to-one synchronous learning environment. *British Journal of Educational Technology*, 42(3):417–440, 2011.

S. Chen, M. Lin, and H. Zhang. Research of mobile learning system based on cloud computing. In *Proceeding of the International Conference on e-Education, Entertainment and e-Management*, 2011.

A. Chickering and Z. Gamson. Applying the seven principles for good practice in higher education (new directions for teaching and learning, no. 47), 1991.

A. W. Chickering and S. C. Ehrmann. Implementing the seven principles: Technology as lever. *AAHE bulletin*, 49:3–6, 1996.

A. W. Chickering and Z. F. Gamson. Seven principles for good practice in undergraduate education. *AAHE bulletin*, 3:7, 1987.

J. P. Chin, V. A. Diehl, and K. L. Norman. Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 213–218. ACM, 1988.

L. Chmiliar. Mobile learning-student perspectives. 2010.

W. C. Chu, Y. Chen, and J.-N. Chen. Context-sensitive content representation for static document. In *Software Engineering Conference, 2004. 11th Asia-Pacific*, pages 719–725. IEEE, 2004.

M. Cieliebak, O. Dürr, and F. Uzdilli. Meta-classifiers easily improve commercial sentiment detection tools. In *LREC*, pages 3100–3104, 2014.

M. Cleveland-Innes and P. Campbell. Emotional presence, learning, and the online learning environment. *The International Review of Research in Open and Distributed Learning*, 13 (4):269–292, 2012.

S. Cloete, C. de Villiers, and S. Roodt. Facebook as an academic tool for ict lecturers. In *Proceedings of the 2009 Annual Conference of the Southern African Computer Lecturers' Association*, pages 16–22. ACM, 2009.

R. S. Cobcroft, S. J. Towers, J. E. Smith, and A. Bruns. Mobile learning in review: Opportunities and challenges for learners, teachers, and institutions. 2006.

J. Cooper and R. Mueck. Student involvement in learning: Cooperative learning and college instruction. *Journal on excellence in college teaching*, 1(1):68–76, 1990.

D. Cormier and G. Siemens. The open course: Through the open door–open courses as research, learning, and engagement. *Educause Review*, 45(4):30, 2010.

J. Costa, T. Ojala, and J. Korhonen. Mobile lecture interaction: Making technology and learning click. In *IADIS International Conference Mobile Learning*, pages 119–124, 2008.

L. J. Cronbach. Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3): 297–334, 1951.

N. Dabbagh, A. D. Benson, A. Denham, R. Joseph, M. Al-Freih, G. Zgheib, H. Fake, and Z. Guo. Massive open online courses. In *Learning Technologies and Globalization*, pages 9–13. Springer, 2016.

K. Dave, S. Lawrence, and D. M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th international conference on World Wide Web*, pages 519–528. ACM, 2003.

D. Davidov, O. Tsur, and A. Rappoport. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd international conference on computational linguistics: posters*, pages 241–249. Association for Computational Linguistics, 2010.

F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340, 1989.

F. D. Davis. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, 38(3):475–487, 1993. ISSN 0020-7373.

F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8):982–1003, 1989.

F. D. Davis Jr. *A technology acceptance model for empirically testing new end-user information systems: Theory and results.* PhD thesis, Massachusetts Institute of Technology, 1986.

G. A. DeBourgh. Use of classroom clickers to promote acquisition of advanced reasoning skills. *Nurse Education in Practice*, 8(2):76–87, 2008. ISSN 1471-5953.

K. J. Denker. Student response systems and facilitating the large lecture basic communication course: Assessing engagement and learning. *Communication Teacher*, 27(1):50–69, 2013.

N. Diakopoulos, M. Naaman, and F. Kivran-Swaine. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *2010 IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 115–122.

H. Du, M. B. Rosson, and J. M. Carroll. Augmenting classroom participation through public digital backchannels. In *Proceedings of the 17th ACM international conference on Supporting group work*, pages 155–164. ACM, 2012.

R. J. Dufresne, W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk. Classtalk: A classroom communication system for active learning. *Journal of computing in higher education*, 7(2): 3–47, 1996.

M. Ebner and M. Schiefner. Microblogging-more than fun. In *Proceedings of IADIS mobile learning conference*, volume 2008, page 155Y159. Algarve Portugal, 2008.

H. F. El-Sofany and S. A. El-Seoud. Towards the development of an m-learning system: A new stage to enhance higher education. *International Journal of Interactive Mobile Technologies*, 3(3):4–9, 2009.

C. K. Eveland, D. A. Socolinsky, and L. B. Wolff. Tracking human faces in infrared video. *Image and Vision Computing*, 21(7):579–590, 2003.

J. Falout, J. Elwood, and M. Hood. Demotivation: Affective states and learning outcomes. *System*, 37(3):403–417, 2009.

S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg. Opinion space: A scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1175–1184, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9.

B. Fasel and J. Luettin. Automatic facial expression analysis: a survey. *Pattern recognition*, 36(1):259–275, 2003.

C. Fellbaum. *WordNet*. Wiley Online Library, 1998.

R. Ferzli and I. Khalife. Mobile cloud computing educational tool for image/video processing algorithms. In *Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE*, pages 529–533. IEEE, 2011.

A. Field. *Discovering statistics using SPSS*. Sage publications, 2009.

C. Fies and J. Marshall. Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, 15(1):101–109, 2006.

J. L. Fleiss, B. Levin, and M. C. Paik. *Statistical methods for rates and proportions (3rd ed.)*. John Wiley & Sons, 2003.

D. Freelon. Recal oir: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, 8(1):10–16, 2013.

E. D. Gagne, C. W. Yekovich, and F. R. Yekovich. *The cognitive psychology of school learning.* Allyn & Bacon, 1993.

M. Gamon. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In *Proceedings of the 20th international conference on Computational Linguistics*, page 841. Association for Computational Linguistics, 2004.

H.-q. Gao and Y.-j. Zhai. System design of cloud computing based on mobile learning. In *Knowledge Acquisition and Modeling (KAM), 2010 3rd International Symposium on*, pages 239–242. IEEE, 2010.

D. R. Garrison, T. Anderson, and W. Archer. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education*, 2(2):87–105, 1999.

V. Gehlen-Baum, A. Pohl, and F. Bry. Assessing backstage - a backchannel for collaborative learning in large classes. In *Interactive Collaborative Learning (ICL), 2011 14th International Conference on*, pages 154–160.

V. Gehlen-Baum, A. Pohl, A. Weinberger, and F. Bry. Backstage–designing a backchannel for large lectures. In *European Conference on Technology Enhanced Learning*, pages 459–464. Springer, 2012.

E. F. Gehringer. Applications for supporting collaboration in the classroom. In *American Society for Engineering Education.* American Society for Engineering Education, 2012.

C. George, E. Bonnie, H. Sarah, and T. Robert. Facilitating online learning: Effective strategies for moderators. *Madison, WI: Atwood Pub*, 2000.

D. George. *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e.* Pearson Education India, 2011.

B. Getzlaf, B. Perry, G. Toffner, K. Lamarche, and M. Edwards. Effective instructor feedback: Perceptions of online graduate students. *Journal of Educators Online*, 6(2):n2, 2009.

M. Gläser-Zikuda, S. Fuß, M. Laukenmann, K. Metz, and C. Randler. Promoting students' emotions and achievement–instructional design and evaluation of the ecole-approach. *Learning and Instruction*, 15(5):481–495, 2005.

A. Go, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1:12, 2009.

C. R. Graham, T. R. Tripp, L. Seawright, and G. Joeckel. Empowering or compelling reluctant participators using audience response systems. *Active Learning in Higher Education*, 8(3): 233–258, 2007.

W. G. Griswold, P. Shanahan, S. W. Brown, R. Boyer, M. Ratto, R. B. Shapiro, and T. M. Truong. Activecampus: experiments in community-oriented ubiquitous computing. *Computer*, 37(10):73–81, 2004.

R. d. Groot. Data mining for tweet sentiment classification. 2012.

J. F. Hair. Multivariate data analysis. 2009.

K. A. Hallgren. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1):23, 2012.

J. T. Hancock, C. Landrigan, and C. Silver. Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 929–932. ACM, 2007.

N. Hara. Student distress in a web-based distance education course. *Information, Communication & Society*, 3(4):557–579, 2000.

A. F. Hayes and K. Krippendorff. Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1):77–89, 2007.

C. Healey and J. Enns. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1170–1188, 2012. ISSN 1077-2626.

T. R. Hensley and M. Oakley. The challenge of the large lecture class: Making it more like a small seminar. *PS: Political Science & Politics*, 31(01):47–51, 1998.

T. J. Hickey and W. T. Tarimo. The affective tutor. *Journal of Computing Sciences in Colleges*, 29(6):50–56, 2014.

A. Holzer, S. Govaerts, J. Ondrus, A. Vozniuk, D. Rigaud, B. Garbinato, and D. Gillet. *SpeakUp A Mobile App Facilitating Audience Interaction*, pages 11–20. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-41175-5.

A. Holzinger, A. Nischelwitzer, and M. Meisenberger. Lifelong-learning support by m-learning: example scenarios. *eLearn*, 2005(11):2, 2005.

K. Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies*, 64(2):79–102, 2006.

K. Hornbæk and E. L.-C. Law. Meta-analysis of correlations among usability measures. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 617–626. ACM, 2007.

J. D. House. Achievement expectancies as predictors of academic performance. *International Journal of Instructional Media*, 20(3):213–23, 1993.

S. Hrastinski. Asynchronous and synchronous e-learning. *Educause quarterly*, 31(4):51–55, 2008.

K.-A. Hwang and C.-H. Yang. Assessment of affective state in distance learning based on image detection by using fuzzy fusion. *Knowledge-Based Systems*, 22(4):256–260, 2009.

W. ISO. 9241-11. ergonomic requirements for office work with visual display terminals (vdts). *The international organization for standardization*, 45, 1998.

S. M. Issack, M. Hosany, and R. Gianeshwar. A me (mobile-elearning) adaptive architecture to support flexible learning. *Malaysian online journal of instructional technology*, 3(1):19–28, 2006.

R. S. Ivanov and M. B. Momchedjikov. Mobile learning and testing with java-enabled phones. *accessed October*, 1(2), 2009.

S. Järvelä, P. Näykki, J. Laru, and T. Luokkanen. Structuring and regulating collaborative learning in higher education with wireless networks and mobile tools. *Educational Technology & Society*, 10(4):71–79, 2007.

A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pages 56–65. ACM, 2007.

D. W. Johnson et al. *Cooperative Learning: Increasing College Faculty Instructional Productivity. ASHE-ERIC Higher Education Report No. 4, 1991.* ERIC, 1991.

R. Junco, G. Heiberger, and E. Loken. The effect of twitter on college student engagement and grades. *Journal of computer assisted learning*, 27(2):119–132, 2011.

A. M. Kaplan and M. Haenlein. The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, 54(2):105–113, 2011.

J. Kirakowski and M. Corbett. Sumi: The software usability measurement inventory. *British journal of educational technology*, 24(3):210–212, 1993.

T. R. Knapp. Focus on psychometrics. coefficient alpha: conceptualizations and anomalies. *Research in Nursing & Health*, 14(6):457–460, 1991.

J. Koeckeritz, J. Malkiewicz, and A. Henderson. The seven principles of good practice: Applications for online education in nursing. *Nurse Educator*, 27(6):283–287, 2002.

B. Kort and R. Reilly. Analytical models of emotions, learning and relationships: towards an affect-sensitive cognitive machine. In *Conference on virtual worlds and simulation (VWSim 2002)*, 2002.

M. Koufaris. Applying the technology acceptance model and flow theory to online consumer behavior. *Information systems research*, 13(2):205–223, 2002.

E. Kouloumpis, T. Wilson, and J. D. Moore. Twitter sentiment analysis: The good the bad and the omg! *Icwsm*, 11(538-541):164, 2011.

U. Krčadinac, J. Jovanović, and V. Devedžić. Visualizing the affective structure of students interaction. In *International Conference on Hybrid Learning*, pages 23–34. Springer, 2012.

U. Krcadinac, P. Pasquier, J. Jovanovic, and V. Devedzic. Synesketch: An open source library for sentence-based emotion recognition. *IEEE Transactions on Affective Computing*, 4(3): 312–325, 2013.

K. Krippendorff. *Content analysis: An introduction to its methodology.* Sage, 2004. ISBN 0761915451.

G. F. Kuder and M. W. Richardson. The theory of the estimation of test reliability. *Psychometrika*, 2(3):151–160, 1937.

D. Kuhn and W. Goh. Arguing on the computer. In *Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*, pages 346–352. International Society of the Learning Sciences, 2005.

A. Kumar and T. M. Sebastian. Sentiment analysis on twitter. *IJCSI International Journal of Computer Science Issues*, 9(4):372, 2012.

T. Kung-Keat and J. Ng. Confused, bored, excited? an emotion based approach to the design of online learning systems. In *7th International Conference on University Learning and Teaching (InCULT 2014) Proceedings*, pages 221–233. Springer, 2016.

J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

P. Lavín-Mera, J. Torrente, P. Moreno-Ger, J. A. V. Pinto, and B. Fernández-Manjón. Mobile game development for multiple devices in education. *iJET*, 4(2):19–26, 2009.

K. Laxman. A study on the adoption of clickers in higher education. *Australasian Journal of Educational Technology*, 27(8):1291–1303, 2011. ISSN 1449-5554.

B. Lent, R. Agrawal, and R. Srikant. Discovering trends in text databases. In *KDD*, volume 97, pages 227–230, 1997.

C. K. Leong, Y. H. Lee, and W. K. Mak. Mining sentiments in sms texts for teaching evaluation. *Expert Systems with Applications*, 39(3):2584–2589, 2012.

J. R. Lewis. Psychometric evaluation of the post-study system usability questionnaire: The pssuq. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 36, pages 1259–1260. SAGE Publications, 1992.

L. S. Lewis, E. T. Pascarella, and P. T. Terenzini. How college affects students: Findings and insights from twenty years of research, 1992.

E. L. Lim and D. W. Moore. Problem solving in geometry: Comparing the effects of non-goal specific instruction and conventional worked examples. *Educational Psychology*, 22(5): 591–612, 2002.

M. Lombard, J. Snyder-Duch, and C. C. Bracken. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4): 587–604, 2002. ISSN 1468-2958.

M. Lombard, J. Snyder-Duch, and C. Bracken. Practical resources for assessing and reporting intercoder reliability in content analysis research projects. retrieved february 24, 2017, 2004.

M. Lombard, J. Snyder-Duch, and C. Bracken. Intercoder reliability. *Retrieved February*, 3: 2011, 2010.

E. MacGeorge, S. Homan, J. Dunning, JohnB, D. Elmore, G. Bodie, E. Evans, S. Khichadia, S. Lichti, B. Feng, and B. Geddes. Student evaluation of audience response technology in large lecture classes. *Educational Technology Research and Development*, 56(2):125–145, 2008. ISSN 1042-1629.

M. Maleko, D. Nandi, M. Hamilton, D. D'Souza, and J. Harland. Facebook versus blackboard for supporting the learning of programming in a fully online course: the changing face of computing education. In *Learning and Teaching in Computing and Engineering (LaTiCE), 2013*, pages 83–89. IEEE, 2013.

M. Marchetti-Bowick and N. Chambers. Learning for microblogs with distant supervision: Political forecasting with twitter. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 603–612. Association for Computational Linguistics, 2012.

A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. Twitinfo: Aggregating and visualizing microblogs for event exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 227–236, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9.

H. W. Marsh. Seeq: A reliable, valid, and useful instrument for collecting students'evaluations of university teaching. *British journal of educational psychology*, 52(1):77–95, 1982.

F. G. Martin. Will massive open online courses change how we teach? *Communications of the ACM*, 55(8):26–28, 2012.

S. Matsumoto, H. Takamura, and M. Okumura. Sentiment classification using word subsequences and dependency sub-trees. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 301–311. Springer, 2005.

D. Maynard, K. Bontcheva, and D. Rout. Challenges in developing opinion mining tools for social media. *Proceedings of the@ NLP can u tag# usergeneratedcontent*, pages 15–22, 2012.

J. P. Mazer, R. E. Murphy, and C. J. Simonds. I'll see you on facebook: The effects of computer-mediated teacher self-disclosure on student motivation, affective learning, and classroom climate. *Communication Education*, 56(1):1–17, 2007.

J. L. McBrien, R. Cheng, and P. Jones. Virtual spaces: Employing a synchronous online classroom to facilitate student engagement in online learning. *The International Review of Research in Open and Distributed Learning*, 10(3), 2009.

D. B. McCabe and M. L. Meuter. A student view of technology in the classroom: Does it enhance the seven principles of good practice in undergraduate education? *Journal of Marketing Education*, page 0273475311410847, 2011.

W. J. McKeachie et al. Teaching and learning in the college classroom. a review of the research literature (1986) and november 1987 supplement. 1987.

R. McSweeney. Sumi–a psychometric approach to software evaluation. *Unpublished MA (Qual) thesis in Applied Psychology, University College Cork, Ireland*, 1992.

P. Melville, W. Gryc, and R. D. Lawrence. Sentiment analysis of blogs by combining lexical knowledge with text classification. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1275–1284. ACM, 2009.

A. Mesbah and A. van Deursen. Migrating multi-page web applications to single-page ajax interfaces. In *11th European Conference on Software Maintenance and Reengineering (CSMR'07)*, pages 181–190. IEEE, 2007.

D. K. Meyer and J. C. Turner. Discovering emotion in classroom motivation research. *Educational psychologist*, 37(2):107–114, 2002.

R. Miller and C. Acton. *SPSS for social scientists*. Palgrave Macmillan, 2009.

A. Minocha and N. Singh. Generating domain specific sentiment lexicons using the web directory. *Advanced Computing*, 3(5):45, 2012.

S. M. Mohammad and P. D. Turney. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465, 2013. ISSN 1467-8640.

N. Morgan and J. Saxton. *Teaching, questioning, and learning.* Taylor & Francis, 1991.

L. F. Motiwalla. Mobile learning: A framework and evaluation. *Computers & education*, 49(3): 581–596, 2007.

S. Motlik. Mobile learning in developing nations. *The International Review of Research in Open and Distributed Learning*, 9(2), 2008.

D. Muijs. *Doing quantitative research in education with SPSS.* Sage, 2010.

T. Mullen and R. Malouf. A preliminary investigation into sentiment analysis of informal political discourse. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 159–162, 2006.

M. Munezero, C. S. Montero, M. Mozgovoy, and E. Sutinen. Exploiting sentiment analysis to track emotions in students' learning diaries. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research*, pages 145–152. ACM, 2013.

A. NCITS. 354: Common industry format for usability test reports. american national standards institute, 2001.

K. A. Neuendorf. *The content analysis guidebook.* Sage publications, 2016.

E. W. Ngai, J. Poon, and Y. Chan. Empirical examination of the adoption of webct using tam. *Computers & education*, 48(2):250–267, 2007.

R. Niazi and Q. H. Mahmoud. Design and development of a device-independent system for mobile learning. *IEEE Multidisciplinary Engineering Education Magazine*, 3(3):63–68, 2008.

D. J. Nicol and D. Macfarlane-Dick. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2):199–218, 2006.

F. Å. Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*, 2011.

K. Nigam, J. Lafferty, and A. McCallum. Using maximum entropy for text classification. In *IJCAI-99 workshop on machine learning for information filtering*, volume 1, pages 61–67, 1999.

G. Noteborn, K. B. Carbonell, A. Dailey-Hebert, and W. Gijselaers. The role of emotions and task significance in virtual education. *The Internet and Higher Education*, 15(3):176–183, 2012.

J. Novak and M. Cowling. The implementation of social networking as a tool for improving student participation in the classroom. In *ISANA International Academy Association Conference Proceedings*, volume 22, pages 1–10. ISANA International Education, Inc Auckland, NZ, 2011.

C. Nunnally Jum and H. Bernstein Ira. Psychometric theory, 1978.

B. O'Connor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM*, 11(122-129):1–2, 2010.

C.-S. Ong, J.-Y. Lai, and Y.-S. Wang. Factors affecting engineers acceptance of asynchronous e-learning systems in high-tech companies. *Information & management*, 41(6):795–804, 2004.

K. Oregan. Emotion and e-learning. *Journal of Asynchronous learning networks*, 7(3):78–92, 2003.

A. Ortigosa, J. M. Martn, and R. M. Carro. Sentiment analysis in facebook and its application to e-learning. *Computers in Human Behavior*, 31:527–541, 2014. ISSN 0747-5632.

A. Pak and P. Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, pages 1320–1326, 2010.

B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135, 2008. ISSN 1554-0669.

B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natu-*

ral language processing-Volume 10, pages 79–86. Association for Computational Linguistics, 2002.

L. Pappano. The year of the mooc. *The New York Times*, 2(12):2012, 2012.

L. Passani and A. Trasatti. Wireless universal resource file (wurfl), 2004.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and d. Duchesnay. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, 12:2825–2830, 2011. ISSN 1532-4435.

R. Pekrun and L. Linnenbrink-Garcia. Academic emotions and student engagement. In *Handbook of research on student engagement*, pages 259–282. Springer, 2012.

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

T. Pelton, L. F. Pelton, and M. Sanseverino. 14. clicker lessons: Assessing and addressing student responses to audience response systems. *Collected essays on learning and teaching*, 1, 2008. ISSN 2368-4526.

S. Petrovic, M. Osborne, and V. Lavrenko. The edinburgh twitter corpus. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, pages 25–26, 2010.

K. A. Pituch and Y.-k. Lee. The influence of system characteristics on e-learning use. *Computers & Education*, 47(2):222–244, 2006.

A. Pohl, V. Gehlen-Baum, and F. Bry. Introducing backstagea digital backchannel for large class lectures. *Interactive Technology and Smart Education*, 8(3):186–200, 2011. ISSN 1741-5659.

A. G. Porter and S. Tousman. Evaluating the effect of interactive audience response systems on the perceived learning experience of nursing students. *The Journal of nursing education*, 49(9):523–527, 2010. ISSN 0148-4834.

A. Poulos and M. J. Mahony. Effectiveness of feedback: The students perspective. *Assessment & Evaluation in Higher Education*, 33(2):143–154, 2008.

D. M. W. Powers. The problem with kappa. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 345–355, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics. ISBN 978-1-937284-19-0.

R. Prabowo and M. Thelwall. Sentiment analysis: A combined approach. *Journal of Informetrics*, 3(2):143–157, 2009.

M. Prince. Does active learning work? a review of the research. *Journal of engineering education*, 93(3):223–231, 2004.

C. Quinn. mlearning: Mobile, wireless, in-your-pocket learning. *LiNE Zine*, 2006, 2000.

D. A. Rahayu, S. Krishnaswamy, O. Alahakoon, and C. Labbe. Rnr: Extracting rationale from online reviews and ratings. In *2010 IEEE International Conference on Data Mining Workshops*, pages 358–368. IEEE, 2010.

M. Ratto, R. B. Shapiro, T. Truong, and W. Griswold. *The Activeclass Project: Experiments in Encouraging Classroom Participation*, volume 2 of *Computer-Supported Collaborative Learning*, book section 57, pages 477–486. Springer Netherlands, 2003. ISBN 978-90-481-6321-2.

K. Regan, A. Evmenova, P. Baker, M. K. Jerome, V. Spencer, H. Lawson, and T. Werner. Experiences of instructors in online learning environments: Identifying and regulating emotions. *The Internet and Higher Education*, 15(3):204–212, 2012.

W. Reinhardt, M. Sievers, J. Magenheim, D. Kundisch, P. Herrmann, M. Beutner, and A. Zoyke. *PINGO: Peer Instruction for Very Large Groups*, volume 7563 of *Lecture Notes in*

Computer Science, book section 51, pages 507–512. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-33262-3.

M. R. Reyes, M. A. Brackett, S. E. Rivers, M. White, and P. Salovey. Classroom emotional climate, student engagement, and academic achievement. *Journal of educational psychology*, 104(3):700, 2012.

M. D. Roblyer, M. McDaniel, M. Webb, J. Herman, and J. V. Witty. Findings on facebook in higher education: A comparison of college faculty and student uses and perceptions of social networking sites. *The Internet and higher education*, 13(3):134–140, 2010.

P. Rodriguez, A. Ortigosa, and R. M. Carro. Extracting emotions from texts in e-learning environments. In *Complex, Intelligent and Software Intensive Systems (CISIS), 2012 Sixth International Conference on*, pages 887–892.

W.-M. Roth, M. K. McGinn, C. Woszczyna, and S. Boutonne. Differential participation during science conversations: The interaction of focal artifacts, social configurations, and physical arrangements. *Journal of the Learning Sciences*, 8(3-4):293–347, 1999.

J. A. Ruiprez-Valiente, P. J. Muoz-Merino, D. Leony, and C. Delgado Kloos. Alas-ka: A learning analytics extension for better understanding the learning process in the khan academy platform. *Computers in Human Behavior*, 47:139–148, 2015. ISSN 0747-5632.

J. A. Russell. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145, 2003.

H. Ryu and D. Parsons. Risky business or sharing the load?–social flow in collaborative mobile learning. *Computers & Education*, 58(2):707–720, 2012.

D. R. Sadler. Formative assessment: Revisiting the territory. *Assessment in education: principles, policy & practice*, 5(1):77–84, 1998.

D. R. Sadler. Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5):535–550, 2010.

R. Sapsford and V. Jupp. *Data collection and analysis*. Sage, 2006.

S. M. Saranya and M. Vijayalakshmi. Interactive mobile live video learning system in cloud environment. In *Recent Trends in Information Technology (ICRTIT), 2011 International Conference on*, pages 673–677. IEEE, 2011.

J. Sauro and J. R. Lewis. Correlations among prototypical usability metrics: evidence for the construct of usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1609–1618. ACM, 2009.

J. Sauro and J. R. Lewis. *Quantifying the user experience: Practical statistics for user research*. Elsevier, 2012.

C. Savill-Smith and P. Kent. *The Use of Palmtop Computers for Learning: A Review of the Literature*. ERIC, 2003.

R. E. Schapire and Y. Singer. Boostexter: A boosting-based system for text categorization. *Machine learning*, 39(2-3):135–168, 2000.

P. A. Schutz and J. T. DeCuir. Inquiry on emotions in education. *Educational Psychologist*, 37(2):125–134, 2002.

E. Scornavacca, S. Huff, and S. Marshall. Mobile phones in the classroom: if you can't beat them, join them. *Communications of the ACM*, 52(4):142–146, 2009.

H. Shachar and S. Sharan. Talking, relating, and achieving: Effects of cooperative learning and whole-class instruction. *Cognition and Instruction*, 12(4):313–353, 1994.

S. K. Sharma and F. L. Kitchens. Web services architecture for m-learning. *Electronic Journal on e-learning*, 2(1):203–216, 2004.

P. J. Shea, A. M. Pickett, and W. E. Pelz. A follow-up investigation of teaching presence in the suny learning network. *Journal of Asynchronous Learning Networks*, 7(2):61–80, 2003.

C. Shi, S. Fu, Q. Chen, and H. Qu. Vismooc: Visualizing video clickstream data from massive open online courses. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 159–166. IEEE, 2015.

K. Siau, H. Sheng, and F.-H. Nah. Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education*, 49(3):398–403, 2006.

R. R. Sims and S. J. Sims. Learning enhancement in higher education. *The importance of learning styles*, pages 1–24, 1995.

J. Sitthiworachart. Interactive mobile learning: Using pdas for higher education in human-computer interaction study. *TECHNOLOGY AND TEACHER EDUCATION ANNUAL*, 18 (4):2105, 2007.

A. A. Skylar. A comparison of asynchronous online text-based lectures and synchronous interactive web conferencing lectures. *Issues in Teacher education*, 18(2):69, 2009.

H.-J. So and T. A. Brush. Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education*, 51(1):318–336, 2008.

A. Sopan, P. J. Rey, B. Butler, and B. Shneiderman. Monitoring academic conferences: Real-time visualization and retrospective analysis of backchannel conversations. In *2012 International Conference on Social Informatics*, pages 62–69.

M. D. Sorcinelli. Research findings on the seven principles. *New directions for teaching and learning*, 1991(47):13–25, 1991.

R. Spence. *Information Visualization: Design for Interaction (2nd Edition)*. Prentice-Hall, Inc., 2007. ISBN 0132065509.

F. K. Stage and D. Hossler. Where is the student? In *Reworking the student departure puzzle*. Vanderbilt University Press, Nashville, 2000.

K. Stephens-Martinez, M. A. Hearst, and A. Fox. Monitoring moocs: which information sources do instructors value? In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 79–88. ACM, 2014.

C. Strapparava and R. Mihalcea. Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, SemEval '07, pages 70–74, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.

D. L. Streiner. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of personality assessment*, 80(1):99–103, 2003.

L. Suen. Teaching epidemiology using webct: Application of the seven principles of good practice. *Journal of Nursing Education*, 44(3):143–146, 2005.

P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen, and D. Yeh. What drives a successful e-learning? an empirical investigation of the critical factors influencing learner satisfaction. *Computers & education*, 50(4):1183–1202, 2008.

M. D. Svinicki. New directions in learning and motivation. *New directions for teaching and learning*, 1999(80):5–27, 1999.

R. Sylwester. How emotions affect learning. *Educational Leadership*, 52:60–60, 1994. ISSN 0013-1784.

B. Szajna. Empirical evaluation of the revised technology acceptance model. *Management science*, 42(1):85–92, 1996.

M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011.

Y.-T. Tai and M.-C. Yang. Integrated platform for collaborative learning in the mobile environment. In *Multimedia and Ubiquitous Engineering, 2008. MUE 2008. International Conference on*, pages 258–262. IEEE, 2008.

Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, 2010.

C. M. Tavani and S. C. Losh. Motivation, self-confidence, and expectations as predictors of the academic performances among our high school students. *Child study journal*, 33(3):141–152, 2003.

J. Teevan, D. Liebling, A. Paradiso, C. Garcia Jurado Suarez, C. von Veh, and D. Gehring. Displaying mobile feedback during a presentation. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, pages 379–382. ACM, 2012.

M. Thelwall. Heart and soul: Sentiment strength detection in the social web with sentistrength. *Proceedings of the CyberEmotions*, pages 1–14, 2013.

M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010. ISSN 1532-2890.

M. Thelwall, K. Buckley, and G. Paltoglou. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173, 2012.

F. Tian, Q. Zheng, R. Zhao, T. Chen, and X. Jia. Can e-learner's emotion be recognized from interactive chinese texts? In *Computer Supported Cooperative Work in Design, 2009. CSCWD 2009. 13th International Conference on*, pages 546–551. IEEE, 2009.

S. Titsworth, M. M. Quinlan, and J. P. Mazer. Emotion in teaching and learning: Development and validation of the classroom emotions scale. *Communication Education*, 59(4):431–452, 2010. ISSN 0363-4523.

J. Traxler. Defining mobile learning. In *IADIS International Conference Mobile Learning*, pages 261–266, 2005.

A. Trifonova and M. Ronchetti. A general architecture for m-learning. 2003.

P. D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417–424. Association for Computational Linguistics, 2002.

T. VanDeGrift, S. A. Wolfman, K. Yasuhara, and R. J. Anderson. Promoting interaction in large classes with a computer-mediated feedback system. *Washington State University*, pages 1–10, 2002.

V. Venkatesh and H. Bala. Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2):273–315, 2008.

V. Venkatesh and F. D. Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2):186–204, 2000.

M. Virvou and E. Alepis. Mobile educational features in authoring tools for personalised tutoring. *Computers & Education*, 44(1):53–68, 2005.

G. C. Visschedijk, A. W. Lazonder, A. van der Hulst, N. Vink, and H. Leemkuil. Modelling human emotions for tactical decision-making games. *British journal of educational technology*, 44(2):197–207, 2013.

K. D. Vohs, R. F. Baumeister, and G. Loewenstein. *Do Emotions Help or Hurt Decisionmaking?: A Hedgefoxian Perspective*. Russell Sage Foundation, 2007.

E. D. Wagner. Enabling mobile learning. *EDUCAUSE review*, 40(3):41, 2005.

C. Ware. *Information visualization: perception for design*. Elsevier, 2012. ISBN 0123814650.

R. R. Weaver and J. Qi. Classroom organization and participation: College students' perceptions. *The Journal of Higher Education*, 76(5):570–601, 2005.

S. Wen and X. Wan. Emotion classification in microblog texts using class sequential rules. In *AAAI*, pages 187–193, 2014.

S. Wendeson, W. F. W. Ahmad, and N. S. Haron. Development of mobile learing tool. In *2010 International Symposium on Information Technology*, volume 1, pages 139–144. IEEE, 2010.

K. Whattananarong. An experiment in the use of mobile phones for testing at king mongkuts institute of technology north bangkok, thailand. *Journal of King Mongkuts Institute of Technology North Bangkok*, 15(1):45–50, 2005.

C. Whissell. The dictionary of affect in language. emotion: theory, research and experience 4. the measurement of emotions, 1989.

C. J. White. Higher education emotions: a scale development exercise. *Higher Education Research & Development*, 32(2):287–299, 2013.

J. White and H. Turner. Smartphone computing in the classroom. *Pervasive Computing, IEEE*, 10(2):82–86, 2011. ISSN 1536-1268.

J. Wiebe, T. Wilson, and C. Cardie. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2):165–210, 2005.

J. M. Wiebe, R. F. Bruce, and T. P. O'Hara. Development and use of a gold-standard data set for subjectivity classifications. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, pages 246–253. Association for Computational Linguistics, 1999.

J. Wu, R. J. Tsai, C. C. Chen, and Y. Wu. An integrative model to predict the continuance use of electronic learning systems: hints for teaching. *International Journal on ELearning*, 5 (2):287, 2006.

Y. Xue, Q. Li, L. Jin, L. Feng, D. A. Clifton, and G. D. Clifford. Detecting adolescent psychological pressures from micro-blog. In *International Conference on Health Information Science*, pages 83–94. Springer, 2014.

C. Yin, B. David, and R. Chalon. Use your mobile computing devices to learn-contextual mobile learning system design and case studies. In *Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on*, pages 440–444. IEEE, 2009.

S. Young and D. G. Shaw. Profiles of effective college and university teachers. *Journal of Higher Education*, pages 670–686, 1999.

H. Yu and V. Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 129–136. Association for Computational Linguistics, 2003.

M. Zembylas. Adult learners emotions in online learning. *Distance Education*, 29(1):71–87, 2008.

L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu. Combining lexicon-based and learning-based methods for twitter sentiment analysis. 2011.

W. Zhao, Y. Sun, and L. Dai. Improving computer basis teaching through mobile communication and cloud computing technology. In *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, volume 1, pages V1–452. IEEE, 2010.

M. Zviran, C. Glezer, and I. Avni. User satisfaction from commercial web sites: The effect of design and use. *Information & Management*, 43(2):157–178, 2006.