Improving Student Engagement Through Visualization of Course Activities

by

Alexandra E. Vermeulen

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Computer Science

Guelph, Ontario, Canada

© Alexandra E. Vermeulen, August, 2014
ABSTRACT

IMPROVING STUDENT ENGAGEMENT THROUGH VISUALIZATION OF COURSE ACTIVITIES

Alexandra E. Vermeulen
University of Guelph, 2014
Advisor: Dr. Judi McCuaig

E-learning is becoming increasingly popular in post-secondary education. However, the engagement levels in an online learning environment remain lower than those in a traditional classroom. This study explores an approach to improving engagement levels in e-learning without the need for extensive student modelling. A Moodle plug-in was created to introduce the aspects of peer competition, comparison, and self progress monitoring. Log data was used to examine any changes in engagement due to the addition of the plug-in. Statistical analysis was inconclusive as to whether or not engagement was improved. However, feedback from the participants was positive and clustering of the data revealed trends that could be incorporated into future iterations of the experiment. The researcher feels that a repeat of either this, or a similar experiment carried out in an experimental environment rather than a classroom may yield better quantitative evaluations.
Acknowledgements

This thesis would never have been completed without the help and support of several wonderful people. First and foremost, I would like to thank my advisor, Judi McCuaig. Without her valuable experience and guidance, I would have had little idea for how to go about completing such a project. With the help of her advice, this thesis turned into something I am proud to have written.

Next i want to thank Scott Dougan, who has been the most dependable partner anyone could ask for. He acted as a sounding board for my ideas on many occasions, was always able to cheer me up when I became frustrated, and was incredibly patient and encouraging throughout the duration of this project. I would also like to thank his parents, Jackie and Richard Dougan for their motivation which helped me to push myself to finish my thesis, as well as Marcia Franz who did a fantastic job editing this document for me.

Last, but most certainly not least, I would like to express my sincerest gratitude towards my parents, Fred Vermeulen and Mary Jo Hind. I would not be here today without their unwavering support, both emotionally and financially. They were always confident that I would complete my Master’s degree even when I was unsure myself.
Contents

1 Introduction 1
  1.1 Thesis Statement 3
  1.2 Overview of Thesis 4

2 Literature Review 5
  2.1 Engagement and Motivation 5
    2.1.1 Motivation Theories based in Social Psychology 7
  2.2 The Importance of Engagement in E-Learning 11
  2.3 Educational Software and Student Modelling 13
  2.4 Detection and Measurement of Student Engagement 17
  2.5 Data Mining and Machine Learning 19
  2.6 Aesthetics and Information Visualization 22
    2.6.1 Aesthetics 23
    2.6.2 Information Visualization 24
  2.7 Increasing Engagement Levels in E-Learning 27

3 Methodology and Implementation 30
  3.1 Proposed Solution 31
  3.2 Materials 32
    3.2.1 Participants 32
3.2.2 Ethical Considerations ........................................ 33
3.2.3 The Course ..................................................... 34
3.2.4 Data Collection and Analysis Resources ................. 36
3.3 Plug-in Design and Implementation .......................... 38
  3.3.1 Objectives .................................................. 39
  3.3.2 Preliminary Design ......................................... 40
  3.3.3 User Interface .............................................. 45
  3.3.4 Implementation ............................................. 50
    3.3.4.i Setting Goals ........................................... 51
    3.3.4.ii Calculation of Class Subsets ......................... 52
    3.3.4.iii Assignment Progress and Overall Progress Graphs 54
3.4 Experimental Design ........................................... 56
3.5 Experimental Procedure ....................................... 58
  3.5.1 Plug-in Deployment ....................................... 58
  3.5.2 Data Collection ........................................... 58
  3.5.3 Participation ............................................... 59
  3.5.4 Data Cleaning and Amalgamation ......................... 61
  3.5.5 Data Analysis ............................................. 65
4 Results and Discussion ................................. 68
  4.1 Summary Statistics of the Data Collected ................. 68
  4.2 Survey Results .............................................. 72
  4.3 Data Exploration ............................................ 76
  4.4 Changes in Variables Indicating Engagement ............ 83
  4.5 Clustering .................................................... 91
  4.6 Discussion .................................................. 102
5 Summary and Future Work ................................. 108
Chapter 1

Introduction

Over the past several years, electronic learning (e-learning) systems have become increasingly popular for use in education. While these systems are hugely beneficial to both students and instructors in many ways, they do have downsides. One major drawback is that students tend to be less engaged and less motivated when learning through an e-learning system than they are with face-to-face instruction (Aragon, Johnson, & Shaik, 2002). A lack of engagement, which often indicates a lack of motivation, has been shown to clearly correlate with a decrease in learning rate (Cocea & Weibelzahl, 2011; Carini, Kuh, & Klein, 2006), making this an important area of research.

There has been some success at improving the level of student engagement during the interaction with educational software. The techniques used tend to fall into one of two camps. One, the approach is all-encompassing, static, and does not target individual students, such as improving the aesthetics and usability of the user interface (UI) or adding multimedia content. Or two, the approach is very customized and dynamically adapts to the students, and therefore requires a detailed student model.
The first approach requires much less work and knowhow from the instructor. The inclusion of multimedia and good aesthetics are important as they have the potential to influence a user’s perception of the usability of a UI (Aljukhadar & Senecal, 2009; Hartmann, Sutcliffe, & De Angeli, 2007). Studies have verified that motivation can be increased simply by creating a more attractive or interactive UI (Barolli, Koyama, Durresi, & Marco, 2006; G.-D. Chen, Shen, Ou, & Liu, 1998). Unfortunately, simply improving the appearance and adding more interactive elements is not enough to bring engagement levels up to the same height as traditional classroom instruction.

The second approach has had success with raising student engagement in e-learning to the levels seen in traditional classroom environments, thus improving the e-learning experience. There have been several studies (Cocea & Weibelzahl, 2011; Tseng, Chu, Hawg, & Tsai, 2008) that measured a significant increase in engagement through analyzing student actions logged by the e-learning system. They achieved this by providing the students with an adaptable and customized learning environment. While this method is certainly more effective in this regard, it is not always practical. In order to provide this level of customization, an extensive student model would be required. Models such as this are not yet supplied with educational software, at least not in a form that would be implementable by the average instructor. If the instructors were able to have such a model created for them, a fair bit of input would be required from the students so that their models can be created. This can be time consuming, unreliable, and something that students would be unwilling to carry out (Villaverde, Godoy, & Amandi, 2006). For these reasons, there is a need for a solution which creates motivation through data that already exists in the majority of e-learning systems.
Several theories in social psychology stress the impact of competition, peer comparison, and social bonds on motivation. Intuitively, it would seem that incorporating these characteristics into educational software would have a positive impact on student engagement. Unfortunately, many of the studies that focus on the topic are based solely on literature rather than experimentation, and therefore remain purely speculation. Researchers who have experimented with incorporating competition, comparison, or group behaviours into educational software have not performed quantitative measurements of their results.

Based on the research in this area, it seems that a more balanced solution is on the horizon. One that is both easily implemented and maintained while also bringing student engagement levels in e-learning to meet those seen in the classroom. By using a simple student model created from existing data, integrating some competition and comparison between students, and creating quality information visualizations for which to present this information, it is probable that engagement can be improved in a manageable way.

### 1.1 Thesis Statement

The thesis statement of this research is that *students who are provided with a visual comparison of their course activities against those of their peers exhibit increased engagement in the class*. In order to demonstrate this, a plug-in for the Moodle e-learning system was created to facilitate visualization of online activities for students and, the interaction patterns of students who used the visualization were analyzed.
1.2 Overview of Thesis

Following this introduction, Chapter 2 begins with a review of the literature on the importance of engagement in the learning process, theories and factors for influencing human motivation and engagement along with methods for integration into an LMS, and data analysis techniques for learning more about how engagement can be improved for students in an e-learning environment. Chapter 3 presents the proposed solution to the problem, the design of the experiment for evaluating that solution, and the procedure for cleaning and analyzing the data collected. The results of data exploration and analysis are discussed in Chapter 4. Finally, Chapter 5 contains a summary of the findings from the study and suggestions for the direction of this work for future research.
Chapter 2

Literature Review

2.1 Engagement and Motivation

For quite some time educators have been aware of the fact that disengaged behaviours have a negative effect on learning. A lack of engagement, which often indicates a lack of motivation, has been shown to clearly correlate with a decrease in learning rate (Cocea & Weibelzahl, 2011; Carini et al., 2006). Because of this, good engagement levels in the classroom are often a priority for instructors, and thus engagement in education is a well-researched subject.

Within the realm of education, the terms engagement and motivation tend to be used interchangeably with one another. Although their meanings are similar, there are some important distinctions between the two labels. Motivation can be defined as “a state of cognitive and emotional arousal which leads to a conscious decision to act, and which gives rise to a period of sustained intellectual and/or physical effort in order to attain a previously set goal” (de Vicente & Pain, 1998). The source of motivation can be either extrinsic, “driven by external rewards or pressure from the environment or from individuals”, or intrinsic, “driven by interest or enjoyment of the activity” (Vassileva, 2012).
Engagement, on the other hand, refers to students’ attentiveness or their willingness to participate in an activity (Ainley, 2004). While motivation is about energy and direction and the reasons for behaviour, engagement describes the connection between the person and the activity (Ainley, 2004). Both can deal with the short term; i.e. across the span of a few minutes or a single learning task; or the long term; i.e. throughout a week or an entire semester. However, it seems that motivation may be harder to measure than engagement since emotion is involved. Therefore engagement is often used as an indicator of a student’s motivation (Cocea & Weibelzahl, 2011, 2007).

In a classroom, it is up to the instructors to engage and motivate their students. Unfortunately, good quality instruction does not necessarily mean that the students will be motivated (de Vicente & Pain, 2002). Some have tried simply adding multimedia material or game elements to their lessons in an attempt to create instruction that is motivating in itself. However, to create engagement, the instruction must be able to adapt to the students themselves. As human beings, we have the innate ability to read each other’s emotions from the observation of body language and actions, and thus, can often tell when someone is becoming disengaged or frustrated. For instance, it is possible to tell that someone is not paying attention if they are staring off into space, fidgeting, or talking to other people. In traditional education, where the teacher interacts face to face with the students, it is fairly straightforward to identify those students that are having trouble staying focused and help them get back on track. Many studies have been done over the years to determine approaches for instructors to create a more engaging environment for their students. Suggestions include things such as demonstrating enthusiasm for the subject, challenging the students, creating collaborative learning experiences, being receptive to student expectations, teaching material the students can
relate to, assessing individual progress, and helping students to expand their self belief (Meece, Anderman, & Anderman, 2006; Said & Al-Homoud, 2004; Zepke & Leach, 2010). The downside to many of these approaches is that they require face-to-face interaction with the students, something that is not always possible, especially with e-learning. What if engagement levels outside of the classroom need to be improved?

One of the most obvious influences of engagement is interest. People tend to be more engaged in activities they are interested in, and thus interest is a determinant of student engagement (Cocea & Weibelzahl, 2007, 2011). Effort is also related to interest in that in the case of uninteresting activities, greater effort is required in order to stay engaged (Cocea & Weibelzahl, 2007, 2011). Research has been done on customizing adaptive learning systems based on personal learning styles and preferences (Tseng et al., 2008). Customization such as this requires the instructor to first obtain information about each student, then apply that information individually to target each student’s styles, preferences, and interests. The incorporation of personalization information into the learning system would certainly be beneficial, however there is a need for a way to do so that is less demanding on the instructor.

2.1.1 Motivation Theories based in Social Psychology

Fortunately, there are other factors that affect human motivation which can be applied to all of the students in the same way. Common bond theory, common identity theory, and social comparison theory all help to explain the formation of a person’s opinions and abilities through social pressures. These theories also give insight to the underlying causes of motivation and how it can be encouraged through the introduction of an appropriate group dynamic.
The *Common identity* and *common bond* theories focus on peers and peer groups of individuals respectively. Both explain how motivation is created through a pull towards uniformity within the group. Either the members want to become more similar to their peers, or they want to help their peers to change in order to make the group more uniform as a whole. The difference between common identity and bond theory is the reason why individuals choose to join and remain part of a group.

Common identity theory explains that individuals join groups because of an identity-based attachment, such as interest in a common topic. Thus, members of identity-based groups are typically attached to the group as a whole or in the theme shared by the group (Grabowicz, Aiello, Eguiluz, & Jaimes, 2013; Raghavun & Vassileva, 2011). Those who participate because of an identity-based attachment may not engage directly with other group members and may even participate anonymously. Because of this characteristic, groups formed from an identity attachment are often seen in online communities.

Common bond states that people join groups based on bond attachment, such as social relationships. Unlike with identity theory, those who join groups based on social bonds are attached to specific members in the group rather than the group as a whole and therefore the main theme of the group may be disregarded (Grabowicz et al., 2013). Bond-based attachments also require members of the group to know each other.

Social comparison theory differs from common bond and identity theory in that it is concerned with the process of comparing oneself to others (Festinger, 1954) rather than the reasons for group formation. This theory describes the tendency of comparing oneself with others in order to evaluate some aspects of the self. Social comparison has been shown to improve task motivation, increase aspiration, and encourage a more skillfull performance (Schunk, 1984).
Festinger (1954) believed that a person’s level of aspiration changes when they are given the opportunity for comparison with others and that, given a range of possible people for comparison, the people closest to one’s own ability or opinion will be chosen for comparison. Likewise, one will be unlikely to evaluate their abilities by comparison with others who are too divergent from oneself (Festinger, 1954; Zanna, Goethals, & Hill, 1974). Suls et al. (2002) later demonstrated that validation does not always derive from comparison with similar others. They found that both upward and downward comparisons can produce positive or negative effects however the variables determining the resulting effect are unclear.

The effect of social comparison is different for those who are close to the mode (the largest subgroup of similar members) than those who are farther away from the mode (Festinger, 1954). Specifically, members who are close to the mode have stronger tendencies to alter the positions of others and weaker tendencies to change their own position compared to those who are distant from the mode. Likewise, those who are distant from the mode have more tendency to alter their own positions rather than change the positions of their peers. In the case of a person who’s performance is considerably higher than that of the group, that person’s performance may decrease slightly over time or they will put in a considerable effort to try to improve the performance of others in the group (Festinger, 1954).

Comparison of one’s own progress can also be beneficial. As people observe their progress, self-efficacy (one’s belief in their ability to succeed) is obtained, which helps to sustain motivation as well as promote skills (Schunk, 1984). The ability of self-assessment has been claimed to be a necessary skill for effective learning (Mitrovic & Martin, 2006). This is also linked to the achievement goal theory, which examines the motivational effect of goal mastery, performance
relative to others and outcomes (Meece et al., 2006). It states that individuals are more likely to engage in a particular task when they expect to do well and when the task has some value to them.

In all three cases, there exists a pressure towards uniformity within the group. In highly cohesive groups (ones that share a strong interest in a similar topic, or have close social relationships) members will attempt to influence each other more or they will be more receptive to influence (Back, 1951). Because of this, the pull towards uniformity is strengthened when the group has a stronger sense of unity or camaraderie. Most often members will help each other because they have the expectation that their help will be compensated or reciprocated by those they have helped, or by the group as a whole (Ren, Kraut, & Kiesler, 2007). This compensation for participation can also come in the form of increased status or reputation within the group, which has shown to clearly provide strong motivation for participation (Vassileva, 2012). There are however people who seem to be immune to reputational incentives, but are willing to contribute to a cause they believe in; either to help their friends (bond-based) or to benefit the community (identity-based) (Vassileva, 2012).

The theories of social comparison, common bond, and common identity have also been applied to studies in the field of education. Raghavun and Vassileva (2011) used social visualizations to motivate participation and reciprocation in online discussions. This was done by presenting the users with visualizations of their reciprocal and non-reciprocal relationships. They found that while awareness was clearly increased, the motivational effect depended on each individual’s pre-existing attachment to the community.

There is a need for more experimentation with applying the theories discussed in electronic learning (or e-learning), where students lack the face-to-face interactions with their instructors and peers. Raghavun and Vassileva (2011)
have applied these theories to increase motivation in online communities, verifying their use for non face-to-face interactions. Since the common identity and social comparison theories do not necessarily require direct encounters between students they could be valuable to bettering engagement in this scenario.

2.2 The Importance of Engagement in E-Learning

Student engagement and motivation have always been a concern to educators. Engagement is linked positively to desirable learning outcomes such as critical thinking and grades and is considered to be among the better predictors of learning and personal development (Carini et al., 2006). In fact, it is so important that the National Survey of Student Engagement (Kuh, 2001) uses engagement levels to measure how students are learning at colleges and universities all across Canada and the USA. Not only is it a significant factor in educational success, the very act of being engaged helps to add to the foundation of skills and dispositions that are required for a productive and satisfying life after college or university (Carini et al., 2006).

With so many aspects learning now being provided by technology; the distribution of learning material, feedback on assignments, discussions between students and so on, it is crucial that the downsides of not having the traditional face to face interactions between students and teachers be considered. One major drawback is that students tend to be less engaged and less motivated when learning through an e-learning system than they are with face-to-face instruction (Aragon et al., 2002). This is a substantial problem since it has been proven that a lack of motivation is clearly correlated with learning rate decrease in e-learning systems where the students do not have face to face contact with the instructor (Cocea & Weibelzahl, 2011; Carini et al., 2006).
Human beings have the ability to read each other’s emotions from the observation of body language and actions, and thus, can often tell when someone is becoming disengaged or frustrated. For example, if someone is staring off into space or talking to someone else, they are unlikely to be giving their full attention. In traditional education, where class sizes are relatively small and the teacher interacts face to face with the students, it is fairly straightforward to identify those students that are having trouble staying focused and help them get back on track. However, many of the cues that help detect other people’s motivation are perceived unconsciously, making it difficult to diagnose motivation without direct observation (de Vicente & Pain, 2003). This is a problem both in e-learning, and where the class sizes become too large for the instructor to observe all of the students.

Improving motivation, and consequently learning, requires an understanding of which motivation-related behaviours most strongly affect the learning process, and then assessing those specific behaviours and motivations (R. S. Baker, Corbett, & Koedinger, 2004). Both educators and researchers now have quite a good awareness of the causes and effects of motivation on learning, as well as how to identify engagement levels in the students. The research explored in the remainder of this literature review yields several technological advancements that help teachers observe the emotional and cognitive states of students remotely. Many of these advancements rely on educational systems’ capabilities to model the students, thus providing the needed information to identify the students’ circumstances.
2.3 Educational Software and Student Modelling

E-learning has become incredibly popular and educational software is now prevalent in most College and University classrooms (Coates, James, & Baldwin, 2005). Web-based Learning Management Systems (LMSs) which have the specific intent of assisting instructors in delivering learning content to students (Machado & Tao, 2007) are most often used. LMSs provide tools for things such as course administration, grade tracking, recording of student activity, and of course, delivery of course material. They can be used to supplement courses taught by instructors in a face-to-face classroom environment, or used exclusively to offer entire courses through distance education.

Because web-based LMSs offer both location and platform independence, allowing students access anywhere with internet and on any device with a web browser, they are usually the most suitable choice for educational institutions (Coates et al., 2005). With a combined market share estimated at 80-90% (in 2007) the two most popular choices for web-based LMSs are Blackboard (paid and closed-source) and Moodle (free and open-source) (Machado & Tao, 2007).

Although e-learning systems can be hugely beneficial to both students and instructors, they do not yet offer the same type of environment and interaction as a traditional face-to-face classroom. The instruction is not as personalized or adaptable, students tend to be less engaged, and there is social and physical isolation among classmates. These drawbacks have a noticeable effect on the learning rate of students (Aragon et al., 2002; Cocca & Weibelzahl, 2011; Carini et al., 2006).

In order to provide a more adaptable and individualized environment, Adaptive Learning Systems (ALSs) were introduced. ALSs attempt to create a more
customized and personalized learning experience by considering things such as the students' learning styles, goals, preferences, knowledge, skills, and interests (Brusilovsky, 2001; Tseng et al., 2008; Weber, 1999). For example, assignment specifications could be presented one way for a sequential learner (an ordered list) and another way for a global learner (overview of the whole picture). Students could also be asked to do different exercises based on their interests, strengths, or weak areas. ALSs function by altering aspects of their interface, functionality and composition based on a student model that exists for each individual (Van Velsen, Dimitrova, Klaassen, & Steehouder, 2008; Akbulut & Cardak, 2012).

These student models, sometimes called learner models, refer to data representations of a student constructed from the observations of interactions between that student and a learning system. They typically contain information about the student’s knowledge, skills, difficulties and misconceptions (Bull, 2004). Some student models also contain information about the learning styles and preferences of the student for closer customization (Kay, 1997; Bull, 2004; Mabbott & Bull, 2004). The models used by adaptive systems are typically hidden from the user. They are used to customize the learning experience for the students, but the students are not able to see what their model contains.

Open learner models (OLMs) are learner models that are made accessible for viewing by the student or other users such as the instructor (Bull & Kay, 2010). OLMs have several advantages over traditional learner models. For both learners and teachers alike, it can be extremely useful to have a visual overview of the students’ activities and how they relate to those of their peers (Duval, 2011). Because the model is observable by the student, OLMs can promote self-reflection and give learners more control and responsibility for organizing their own learning (Hansen & McCalla, 2003; Bull & Kay, 2010; Bull &
Learners also can become aware of their current knowledge, and in some cases that of their peers, allowing them to more easily identify goals (Kay, 1997). These aspects of OLMs are valuable because they support the acquisition of self-knowledge, which is an important skill that leads to self-directed life-long learning (Kay, 1997). On a more short-term basis, studies have demonstrated that OLMs can improve the performance of less able students while also boosting the self-confidence of the more able students, a considerable benefit in distance learning environments (Mitrovic, 2007).

OLMs have also shown to be valuable in lessening the feeling of social and physical isolation between students. In some cases, they have been implemented in online learning to give the students a better sense of their classroom community (Dufresne & Hudon, 2002). Students are able to compare their progress to others, they have more awareness about what is happening in the group, and can more easily compare and discuss information amongst themselves. In fact, Duval (2011) has observed that the strength of influences from online social networks appear greater than those in real-world social networks. If so, educators could use OLMs to harness the power of social links between students in order to promote collaborative and competitive learning.

There are several forms of OLMs that are commonly implemented in e-learning environments. Depending on the information that is being shared and the type of interaction desired with the students, different variations may be used. Bull and Kay (2010) classify these variations into three types of learner models:

1. *Inspectable* models are viewable by the student but remain completely under control of the system, not allowing the student to make or suggest changes.

2. *Editable* models can be viewed as well as altered by the student.
3. *Challenge* models are an amalgamation of the inspectable and editable versions. The student can make changes, but only if they justify the changes to the system, adding further information to the model in the process.

Customization of OLMs may also involve changing what information is being presented and how that information is visualized. Over the last 15 years there has been a fair amount of debate over what kind of information the model should provide (Self, 1994). Although the contents are often chosen to fulfill a certain goal, more recent OLMs commonly include information about the students’ progress within a course, their knowledge level, and some sort of expected knowledge level for comparison.

Research has indicated that when students observe their progress, self-efficacy is obtained which helps to sustain motivation and promote skills (Schunk, 1984). OLMs also help to support self-assessment, which has been deemed necessary for effective learning (Mitrovic & Martin, 2006). In a study by Mitrovic (2006) OLMs were used to improve the performance of the less-able students while also boosting the confidence of the more-able students. This was done by simply providing the students with visualizations that showed them their own proficiency (percent covered and percent learned) in each section of the course. OLMs support self-assessment (Mitrovic & Martin, 2006), create a more individualized learning experience (Bull & Nghiem, 2002), and reduce isolation between students (Dufresne & Hudon, 2002), thus amending three disadvantages encountered with e-learning.

E-learning software is a valuable asset for research on how students learn. The majority of e-learning programs have the ability to record student actions and information in the form of log files. Mining and analysis of such data can help to identify areas where students struggle, link success with certain student
behaviours, and evaluate the introduction of new features in e-learning systems. Using this log data researchers have been able to detect and measure student engagement levels, something that is typically done through direct observation of the students.

2.4 Detection and Measurement of Student Engagement

With so many courses utilizing e-learning systems or large class sizes, many students lack direct interactions with their instructors. One major downfall of this is that the instructor is not able to identify when students become disengaged and then instinctively get them back on track. Previous attempts to improve online learning systems have focused on motivational planning, such as incorporating more attractive designs and multimedia materials to motivate students (G.-D. Chen et al., 1998), and have ignored motivation diagnosis or measurement (de Vicente & Pain, 1998). As previously discussed, motivated students learn much more effectively than unmotivated ones, therefore it would be beneficial for the system to be able to identify disengaged behaviour and act accordingly. Recent research has demonstrated that engagement can be diagnosed by examining various attributes recorded by the e-learning system using machine learning techniques (Beck, 2004; Qu & Johnson, 2005).

In the past, researchers have used techniques such as questionnaires, self-reports, verbal communication, and expert observation to elicit motivation and engagement information from students (de Vicente & Pain, 1998). However, each of these methods has limitations. If students’ learning process is being interrupted to ask for a self-report on their current motivation, it is likely that the interruption itself is dissipating whatever motivation they had. The use of verbal communication and expert observation require the students to be observable, which they are often not when considering online learning systems. Lastly,
questionnaires have been deemed often unreliable, and a time consuming task that students are not always willing to carry out (Garcia, Amandi, Schiaffino, & Campo, 2005; Villaverde et al., 2006; Yannibelli, Godoy, & Amandi, 2006). Regardless, the survey approach is still widely used for its ease of administration. Handelsman (2005) developed and explored the validity of the Student Course Engagement Questionnaire (SCEQ) which measures absolute engagement of students in their present course. This survey appears frequently in the evaluation of University and College courses and, relative to similar surveys, is said to be quite reliable, valid, and multidimensional.

De Vicente and Pain (2003) demonstrated that the motivational state of a student can be identified without being able to observe the students themselves. Their results showed that it is possible to infer motivation diagnosis knowledge based on the information provided by log files of the computer interaction within an e-learning system. Similarly, in 2004 Baker et al. (2004) were able to automatically identify students who were taking advantage of the system’s help feature to get answers without first attempting the questions. Then in a recent study they expanded on this research and developed a set of detectors that were somewhat successful at identifying engaged concentration, confusion, frustration, and boredom in students (R. S. J. d. Baker et al., 2012). Also in 2004 Beck (2004) found that response times from log files can be used to model student engagement. Shortly after, Qu and Johnson (2005) demonstrated that confidence, confusion and effort could be inferred partially from the expected time to perform a task. Cocea and Weibelzahl (2006) expanded on this success from looking at activity tracking log files and developed a decision tree for assessing users’ motivation. Their methodology was also verified to work with other systems and learning environments (Cocea & Weibelzahl, 2007).
Although the studies mentioned vary in their techniques for measuring engagement levels, there are many similarities between which log file attributes were examined. Variables commonly used included:

- Start and end times, time on task, and response times of various actions or modules
- Login frequency and time between logins to the e-learning system
- Number of lessons/assignments/quizzes read and completed (also attendance in courses that also have in-class lectures)
- Grades and changes in grade
- Forum participation

The results from the experiments discussed above have shown that engagement tracking is not only possible, but could likely make its way into e-learning systems in the not-so-distant future. This would allow for more definitive evaluations of whether or not a new feature helps or hinders student engagement. The variables stated above are already logged by most e-learning systems, and little work is required to extract and transform the data into a workable format. By examining variables such as this one can measure the changes in engagement as a result of developments in the e-learning system, thus ensuring that the system is only made better in this regard.

2.5 Data Mining and Machine Learning

The examination of interaction patterns and learning behaviours of students can help to identify ways in which educational software can be improved. E-learning systems accumulate an extensive amount of information that is valuable for analyzing these attributes (Mostow & Beck, 2006). The use of data mining and machine learning on the information collected has brought about the discovery of much new information on the ways students learn.
In general, data mining refers to the process of analyzing data from various perspectives and summarizing it into useful information (Palace, 1996). Educational data mining (EDM) is a field that uses data mining and machine learning techniques on educational data specifically with the goal of better understanding the learning process (Romero & Ventura, 2010). This aids in identifying what changes help the students and which ones hinder them (Mostow & Beck, 2006).

Machine learning is a branch of artificial intelligence based on mathematical algorithms and automation. Analytic models are created that use algorithms to perform tasks by generalizing from examples (Domingos, 2012). The algorithms used are often genetic algorithms which borrow their design from nature and are designed to work on large spaces involving states that can be represented as computational strings (Goldberg & Holland, 1988). With each subsequent iteration of the algorithm, the “machine” learns from its mistakes in previous steps to derive the best results without requiring human intervention. There are many classes of machine learning algorithms, however it seems classification and clustering are the most used within the realm of EDM.

Classification algorithms tackle the problem of identifying in which class or category an observation belongs. The categories (or labels) are provided, and the classifier is trained to apply the proper label using an example data set. Classification is considered supervised learning, meaning that the “machine” is trained on a set of correctly labeled observations (Alpaydin, 2010). Because of this, it is useful for prediction or to label new observations when the possible outcomes are known. In e-learning research, classification has been used with data collected via educational software to discover potential student groups with similar characteristics (G. Chen, Liu, Ou, & Liu, 2000), predict performance and final grades (López, Luna, & Ventura, 2012; Minaei-Bidgoli
& Punch, 2003), recognize learning styles (Chang, Kao, Chu, & Chiu, 2009), and identify students with low motivation (Cocea & Weibelzahl, 2006).

Clustering or cluster analysis involves grouping data into categories based on some measure of inherent similarity (Alpaydin, 2010). Labels are not provided for clustering algorithms, only the data set is used for input. For some implementations the desired number of clusters also needs to be specified, others determine the optimal number of clusters automatically. Unlike classification, it is considered to be an implementation of unsupervised learning since it attempts to find the underlying structure in unlabelled data (Alpaydin, 2010). This characteristic makes clustering more useful for discovering unknown qualities in the data. It cannot be used for prediction or to label single observations since the output is a grouping rather than classification and there is no specified target attribute. In EDM clustering has been used to find groups of students with similar behaviour patterns (Bovo, Sanchez, Hégyi, & Duthen, 2013b) and to link the causality of attributes (such as how student participation impacts final grades (López et al., 2012).

There are several clustering algorithms available, however previous studies in EDM have reported the most success with the Simple K-means algorithm (Bovo et al., 2013b; Romero, Ventura, & García, 2008) and the Expectation-Maximization (EM) algorithm (López et al., 2012; Slaninová, Martinović, Drážilová, & Snašel, 2014). Besides choosing an algorithm, one must also determine the best number of clusters to use. Bovo (2013a) states that most clustering algorithms (EM, K-Means, X-Means and Hierarchial Clustering were tested) generally agree on at most two or three clusters. One cluster is useless, since no divisions into groups are made, and more than three clusters tended to result in outlier groups that contained too few samples to be helpful for describing the data.
Because there is such a range of classification and clustering algorithms available, a machine learning software suite (an application containing a collection of data mining algorithms) may be beneficial. The Weka (Hall et al., 2009) system comes highly recommended (Bovo et al., 2013b; Romero et al., 2008). Weka is free, implemented in Java, and contains a selection of machine learning algorithms so that one can easily experiment with several in order to achieve the best results.

Acquiring and analyzing the data are only the first steps in presenting new results. The way in which data is presented visually has a huge impact on how the information is interpreted. It is easier for people to see patterns in the data through graphical representations than through a simple chart of data (Ware, 2013, p. 213). Both the aesthetics and the design of the visualization itself need to be considered in order to properly represent the intended results.

2.6 Aesthetics and Information Visualization

Norman (2004, p. 81) states that there are four components of good design: “Function, understandability, usability, and physical feel.” While the first three components are the most important objectively, the fourth can completely change a person’s subjective opinion of the other requirements. If a product doesn’t function, isn’t understood, or is unusable then it is useless. (Consider a knife that doesn’t cut, or a microwave that no one can figure out how to turn on.) On the other hand, a product can be completely functional but lack any aesthetic appeal (a sense of beauty or taste). Utility and usability are important, but without any fun, pleasure, joy, or excitement people are unlikely to use the product (Norman, 2004, p. 19).
The quality of information visualizations and overall aesthetic appeal are important considerations when designing a UI. Research indicates that the design and aesthetics of software or products contribute to the motivation and engagement of the users (Barolli et al., 2006; G.-D. Chen et al., 1998). The incorporation of multimedia elements, which includes replacing textual information with information visualizations, helps to increase perceived interactivity and interest (Aljukhadar & Senecal, 2009). While the improvement of aesthetics alone does not create a big enough change in engagement to be a sole solution, it is certainly something that should not be overlooked.

2.6.1 Aesthetics

Aesthetic appeal (one’s attraction to the visual aspects of something) is an important aspect of UI design. It has been shown that perceived usability is higher for appealing products than for unappealing ones, even if the devices are identical in terms of objective usability (Sonderegger & Sauer, 2010). In fact, some claim that aesthetic design can be a more important influence on user preference than actual usability (Hartmann et al., 2007; Norman, 2004, p. 17). There have also been correlations between the perceived aesthetic quality of an interface and overall user satisfaction (Lindgaard & Dudek, 2003). Other studies have verified that simply creating a more attractive interface can increase the motivation of students (Barolli et al., 2006; G.-D. Chen et al., 1998). While the improvement of aesthetics alone does not create a big enough change in engagement to be a sole solution, it is certainly something that should not be disregarded.

When planning the design of an interface, focus should be put on the intuitiveness and ease-of-use (Aljukhadar & Senecal, 2009). The interface can be made dynamic so as to present the user with options which are relevant to
the task at hand, therefore simplifying and streamlining the user’s work flow. The look of the interface should be pleasing as well, so as to create good subjective as well as objective usability (Norman, 2004, p. 19). Choices such as the appearance and position of buttons, the order of menu items, the layout of content, and the selection of colours all have an impact on how users perceive an interface.

Colour has an impact on how a person feels. Cooler colours, such as blues and greens, are typically more favoured than warmer colours, such as yellows and reds (Cyr, Head, & Larios, 2009). By carefully selecting a colour palette, one can shape the user’s emotional response. For instance, the colour blue is calming and creates a sense of trust, orange denotes cheapness, and red increases one’s heart rate, therefore putting them in a more alert state (Cyr et al., 2009; World & partners, 2014). Because of these properties, colours should be suited for the context and culture in which they are used so as to make the interface intuitive (for example, green to indicate something good and red to alert the user to something bad).

2.6.2 Information Visualization

Information visualization, or simply visualization within the context of this thesis, refers to a graphical representation of data or concepts with the aim of maximizing comprehension by presenting the information intuitively and naturally (C. Chen, 2006, p. 26, Romero, Ventura, Pechenizkiy, & Baker, 2011, p. 10, Ware, 2013, p. xvi). It is now a prevalent trend in modern UI design (C. Chen, 2006, p. 1). The inclusion of visualizations is thought to enhance a student’s learning by presenting data in a way that is more interesting (Naps et al., 2003). However, there appears to be a disconnect between this belief and the willingness for instructors to deploy such visualizations in their classroom (Naps et
al., 2003). By using visualizations for the presentation of an OLM (refer to Section 2.3), the instructor is not required to create the graphics and the students are provided with a customized visual representation of the information that is pertinent to each of them. Visualizations used for this purpose would need to be designed so that they are immediately understood without training.

According to Ware (2013, p. 3), “One of the biggest benefits of data visualization is the sheer quantity of information that can be rapidly interpreted if the data is presented well”. He also states several other advantages. The use of visualizations allows for the recognition of properties that were not expected. Likewise, problems with the data may become apparent. Visualizations can also facilitate both the understanding of features of the data and the formation of hypotheses.

The decision of what strategies and tools to use for a particular type of visualization is a challenging one (C. Chen, 2006, p. 1). It is important to have a clear goal for the intent of the visualization, and a plan for how to evaluate whether that goal has been achieved (Duval, 2011). A large part of the assessment must look at how well a user can use, or understand, the graphic. Freitas et al. (2002) evaluates the usability of information visualizations in four distinct areas:

- Completeness (representative of all meaningful contents of the data)
- Spacial organization (overall layout of the visual representation)
- Codification of information (use of intuitive symbols or realistic characteristics to aid in understanding of the information)
- Changes in spacial organization (rebuilding after a user action and the time to do so)
The size of a graph, which affects spacial organization, is a common issue in graph visualization. Comprehension and analysis of data in a graphical format is easiest when the size of the visualization is small (Herman, Melancon, & Marshall, 2002). Herman et al. (2002) stresses the point that visualizations must be created in real-time so that the user always sees the current information. This would be especially important when displaying student model data in an E-learning environment. Herman also advises the use of 2-dimensional figures as opposed to 3-dimensional ones as objects in 3-D can obstruct one another.

With regard to visualizations, researchers have supported both static presentation (the same for every student) as well as representations that are adapted to each individual student. Mabbott and Bull (2004) for example, feel that the learners’ models should be presented in whichever form suits them best, considering aspects such as learning style for the customization. For instance, perhaps one student prefers to see his or her progress as a graph, and another prefers to see it as a list of which items have been completed. The motivation behind this thought is that reflection may be increased if the learners can view their model in the form they prefer. Contrary to this opinion, Vassileva and Sun (2008) state quite clearly that customization should be avoided. They feel that “all users should see the same thing so that they feel responsible for their actions, since they know that others see the same things as them and are aware of what they do” (Vassileva & Sun, 2008). While both approaches have their advantages, the effectiveness of the static model is more straightforward to evaluate in an experimental setting.
2.7 Increasing Engagement Levels in E-Learning

The literature explored in this chapter outlines much of what researchers have discovered about measuring, detecting, and improving human motivation. With this knowledge, it seems very likely that motivation and engagement could be improved within e-learning systems, therefore creating a better learning environment for the students. Some methods for increasing engagement have been thoroughly studied within the realm of e-learning software. However, the approaches suggested by social psychologists are not yet prevalent in the context of e-learning.

The research on the impact of aesthetics and interactive interface design concludes that these are important aspects to an e-learning system (see Section 2.6). Aesthetics have been shown to increase engagement (Barolli et al., 2006; G.-D. Chen et al., 1998), but not enough to be a sole solution. They do however have a very strong influence on user satisfaction and the perceived usability of an interface (Aljukhadar & Senecal, 2009; Sonderegger & Sauer, 2010). If something has a poor visual design, one automatically tends to assume it functions poorly as well. This first impression tends to remain even after using and becoming familiar with the product (Sonderegger & Sauer, 2010). It would be extremely unfortunate if students were unmotivated simply because they were unsatisfied with the aesthetics of the learning system. Because the use of e-learning is usually mandatory for the students, and their aesthetic preference could impact their long-term satisfaction and performance, the look and feel of the software should always be considered.

On the other hand, researchers have been able to increase engagement substantially by providing students with a very customized and adaptable learning environment (Cocea & Weibelzahl, 2011; Tseng et al., 2008). This method is intended to improve motivation by targeting students’ interests, learning styles
and other preferences. If the implementation uses an OLM then learners can also be reminded of their knowledge gaps and progress (according to the system), allowing them to more easily identify and pursue their goals (Kay, 1997). Unfortunately, this approach requires extensive student modelling (as discussed in Section 2.3) in order to achieve such a high level of customization. While it may be worth the effort in some situations, the requirements at this time are often impractical for the average class and instructor.

Social psychology provides an opportunity for a third approach. Social factors, as supported by several of the theories introduced in Section 2.1 can also play a large role in motivating people. This includes the comparison, competition, and camaraderie introduced by a group dynamic. Researchers in the area of OLMs have observed the tendency of students to compare and compete with their peers (Bull & Britland, 2007; Mabbott & Bull, 2004; Bull & Nghiem, 2002; Linton & Schaefer, 2000). In earlier experiments, examining individual models, the students expressed an interest in sharing models, showing that compiling a group model or allowing for comparison may be beneficial (Mabbott & Bull, 2004; Bull & Nghiem, 2002). A study by Bull and Britland (2007) gave students the option to share their own learner models with their peers, facilitating comparison with their peers. The students were generally more interested in the models of those they considered stronger than themselves or those who had a similar ability level, as predicted by Festinger’s (1954) social comparison theory. Some students in this study even made statements such as “I became more competitive. I no longer just wanted to meet the weekly target. I wanted to race ahead of the weekly target and also my peers.” The very positive feedback from this study show that having access to the models helped to motivate the students to get ahead in the course. However, analysis
was based only on student feedback so there was no quantitative measurement that verified such an effect.

Studies which incorporate peer competition and comparison into e-learning appear to be few and far between. It appears that the next logical step is to introduce these social factors of competition, comparison, and progress monitoring into educational software, then quantitatively measure the effect on engagement through data mining techniques. Previous studies have identified which attributes best detect engagement levels, and most e-learning software already logs these variables. OLMs can be utilized to generate information visualizations that display the individual students’ progress as well as the progress of their peers. The models can be generated at the time they are viewed, guaranteeing the most up-to-date representation. Favourable measurements of engagement due to the introduction of these features could mean a big step forward for e-learning.
Chapter 3

Methodology and Implementation

As stated in the introduction (Section 1.1), the thesis of this research is that students who are provided with a visual comparison of their course activities against those of their peers exhibit increased engagement in the class. This chapter describes an experiment that explored such an approach for increasing student motivation within an e-learning environment through the use of personalized and comparative visualizations of their progress within a class.

Moodle, which stands for “Modular Object Oriented Developmental Learning Environment” is an open-source and web-based Learning Management System. Owing to the fact that it is open source, Moodle allows for a wide range of user customization and the ability to create third party plug-ins or modules. For this study a Moodle plug-in, referred to as the Progress Tracker, was created to display both class and individual progress overall and on assignments as well as provide a tool for goal setting and monitoring. This plug-in was made available to students in a second year programming class for the latter portion of a semester. Data for the study was captured both before and after the
addition of the plug-in and was analyzed to determine if student engagement increased in the duration the plug-in had been available.

3.1 Proposed Solution

In Chapter 2, the literature review discusses several shortcomings of the existing methods for improving student engagement. Many of the current solutions are either generic, and therefore do not target the individual student (such as improving aesthetics or adding multimedia), or alternatively, require an extensive student knowledge model for customization.

With the merits and downfalls of the existing approaches in mind, the researcher of this thesis proposes a solution that combines aesthetics, interactivity, individualized student progress monitoring, goal setting, and peer competition and comparison. The goal is to increase engagement with a simple model that utilizes information already existing in the majority of e-learning systems, therefore requiring little maintenance and effort on the part of the instructor.

This approach presents a model that provides students with graphical visualizations of their progress on assignments and the course overall. Students can set goals for their grade in the course, and are able to see their progress towards their goal each time they log into the e-learning system. They can also see where they stand compared to their classmates; both in the course and on the current assignment. This information was presented using an inspectable learner model. The students can see their progress and their classmates’ progress in real-time, but they cannot make or suggest changes to the contents.
3.2 Materials

3.2.1 Participants

Participants for this study were recruited from the Winter 2013 offering of the CIS*2500 Intermediate Programming course at the University of Guelph (see details in Section 3.2.3). CIS*2500 is a first year, second semester course primarily comprised of Computer Science and Engineering students.

This particular class was chosen for several reasons. CIS*2500 has a mandatory prerequisite of the Introduction to Programming course, CIS*1500. As such, the students were familiar with the Moodle software and had previous experience with computer programming, and thus, the expected time commitment for programming assignments. The class size was also fairly large at approximately 200 students. Because the Progress Tracker plug-in was only being incorporated into a single class, it was important to have enough students to recruit for the data collection phase and maintain a decent sample size (preferably over 50 samples). Finally, CIS*2500 was being taught by the supervisor for this research. Because she was the administrator for the course Moodle page, it was convenient for the plug-in to be incorporated into her class.

Data were collected from the 75 participants that consented to be part of the study. Unfortunately this number was lower than expected due to the limited time in which recruitment was carried out. As an incentive to participate, all participants were entered in a draw for one of two $25 gift certificates to Future Shop.
3.2.2 Ethical Considerations

This study was approved by the Research Ethics Board (REB) of the University of Guelph. Documents containing a much more detailed description of the experiment and a record of approval are available from the REB file number 13MR006.

Students from the CIS*2500 class were informed of the study during the last week of classes (the REB did not grant approval until this point). Data had been logged for the entire semester, however consent was required for the analysis of each participant’s data. It was made clear that their participation, or lack thereof, would have no impact on their grades in the course. All students in CIS*2500 were invited to participate. Consent forms were handed out and collected for those wishing to be part of the study. The students also were told that, should they change their mind about participating, they could withdraw from the study at any point before the participants’ identification was removed from the data (estimated to be May 2013). Neither the data nor consent forms were made available to the researcher until after final grades had been released by the university.

Data for the students could not be anonymous. For data analysis it was required that all of a participant’s data remain linked to that participant, however any information that would link a student’s identification with a participant was discarded.

In all cases, identifying information was removed from the data and was replaced by participant ID numbers during data cleaning. This included names, student IDs, student numbers and email addresses. Additional precautions were taken to reduce the likelihood of being able to backtrace the data and
link a participant to a student. These precautions included rounding any timestamps to the nearest hour and converting all grades to the letter grade scale used by the University of Guelph (see Table A.2 in Appendix C). This scale provided 13 bins for grades to be sorted into, therefore maintaining what the researcher considered to be a reasonable level of accuracy for all grades except for “F” which covered everything from 0% to 49%.

3.2.3 The Course

CIS*2500 is an intermediate level course in C programming. Students are required to have successfully completed the prerequisite programming course, CIS*1500, before taking it. By the end of CIS*2500 it is expected that students can read and interpret a program specification and implement it as reliable code.

The course was structured using blended learning, meaning that the students were taught using both e-learning and traditional classroom instruction. The in-class component consisted of three fifty-minute lectures each week over a 12 week semester. Students were also expected to attend one lab session per week. Data were collected from the first day of class (January 6, 2013) until official grades were released (April 25, 2013).

Grades for CIS*2500 weighted slightly more towards coding components than written ones. The coding portion of the course counted for 55% of the final grade. For this portion students completed three assignments, each worth 10% of their grade. The assignments were submitted using Git, a distributed revision control and source code management system (Chacon, 2009). Students also had three lab exams, held during their weekly lab sessions, which were worth a total of 25% towards their grade (5%, 10% and 10% respectively.)
The written portion was worth 45% of the final grade. It was made up of two quizzes which were held during lectures and counted for 15% of the grade (5% for the first quiz and 10% for the second.) Students in CIS*2500 were required to bring wireless devices called *clickers* to class. The clickers were used to respond to multiple choice questions asked during the lectures. Responses were recorded and students were given a participation mark based on their clicker usage. This mark was worth 10% of their grade. The final 20% of the grade was allotted to the final exam for the course. Students were required to pass the exam in order to pass the course, otherwise their grade was recorded as the minimum between their exam grade and calculated grade.

The School of Computer Science at the University of Guelph has been using Moodle for several years to present online components for many of its classes. The installation of Moodle referred to in this thesis is the one used for the CIS*2500 course website. This instance was hosted on Dr. Judi McCuaig’s server and uses a Postgres database for its backend.

Git is a free and open source distributed version control and source code management system, similar to SVN which is also commonly used (Chacon, 2009). The use of version control is essential in real-world developing, particularly when working with teams of programmers. In order to encourage proper development practices, students in CIS*2500 were required to use Git to submit the code for their assignments. Students were also expected to regularly record changes in their assignments to their Git repository via *Git commits*. The last Git commit before the due date of the assignment was the assignment used for submission. Although preferred, they were not marked on whether they used Git regularly throughout their completion of each assignment.
As proposed in Section 3.1, the Progress Tracker plug-in was created with the aim of increasing student engagement through peer comparison and competition. The plug-in was implemented as a Moodle module and deployed for use by the students of CIS*2500 during their final assignment. (Refer to Section 3.3 for details of the design and implementation.)

3.2.4 Data Collection and Analysis Resources

Data for the study were acquired from five different sources and aggregated to create one main data set.

1. Goal setting information from the Progress Tracker plug-in was obtained from the Moodle database.
2. Activity logs and responses to questionnaires (detailed in Section 3.4) were downloaded from the Moodle website itself.
3. Grades were provided by the course instructor since her copy contained any last minute changes made before submitting the official grades to the University.
4. Assignment submission times were gathered from the Git commit logs of each participant.
5. Attendance information was collected from recorded in-class question participation.

The data collected were analyzed using R (R Core Team, 2014), a software environment for statistical computation, and Weka (Hall et al., 2009), a collection of machine learning and data mining algorithms. R was used to find correlations and statistical significance, to display likert data from questionnaires, and to create various other charts and graphs illustrating the relationships in the data. Weka is open source software that provides a collection of pre-programmed machine learning and data mining algorithms (Witten & Frank,
2005). For this study, Weka was used for data pre-processing, classification, clustering, and visualization.

Decision trees, neural networks, and clustering were all considered for analysis of the data. Several decision tree algorithms were used with the data, yet none were able to create a meaningful tree (in this case more than 2 nodes) or a tree with an accuracy much higher than chance. Neural networks were also explored, but resulted in very high error rates, even when the data was grouped into only 2 or 3 bins per attribute. For these reasons decision trees and neural networks were rejected as forms of analysis on this data in favour of clustering.

The K-means clustering algorithm was used for the clustering of this data. This algorithm was chosen for its demonstrated success with e-learning data of similar types in other studies (Bovo et al., 2013b; Romero et al., 2008). Exploration was also done with the Expectation Maximization (EM) clustering algorithm which has also shown success in clustering students (López et al., 2012; Slaninová et al., 2014), however the K-means algorithm resulted in much clearer distinctions between groups for the data in this study. Weka allows for customization of several parameters of this algorithm. The available options are listed (and set to the option used in the study) in the screenshot of the algorithm customization menu from Weka shown in figure 3.1.

Display standard deviations was set to true in order to view the distribution of category type attributes. Without this option only the category with the highest count in each cluster is displayed. Missing values were not replaced since the researcher was interested in knowing when participants opted out of completing components. The number of clusters was set to three after comparing the results of using 2 to 5 clusters. Three clusters provided groups of a large enough size, yet with enough segregation to see the differences between
participants assigned to each cluster. A detailed discussion of this testing is included in Section 4.5. The remaining parameters were left on their default settings as shown in figure 3.1.

![K-means clustering algorithm parameters as seen in Weka](image)

Figure 3.1: K-means clustering algorithm parameters as seen in Weka

### 3.3 Plug-in Design and Implementation

The goal of the plug-in was to increase student engagement through a combination of self-progress information, goal setting, and peer comparison. This information was displayed primarily through data visualizations. The decision of what information to present was made through brainstorming with the use of wireframes (details of prototypes follow in Section 3.3) in combination with research on factors increasing human motivation (discussed in Section 2.1). The remainder of this section describes the design and implementation process in detail.
3.3.1 Objectives

The objectives for the plug-in were formed around motivation inducing methods discussed in the literature review (Chapter 2). Goal setting, progress monitoring, and competition have all been used successfully to improve engagement and motivation in a variety of environments. Therefore, the objectives of the plug-in were as follows:

1. To provide students with a tool for goal-setting that also keeps them aware of their progress towards that goal
2. To remind students of their progress on course content and how much time remained before the due date of that content
3. To allow students to compare their progress to that of their peers, and possibly introduce a competitive aspect into the e-learning system

In order to fulfill the goal-setting objective, the plug-in was created so that students could assign a goal grade to themselves for the class. Once a goal was set the students were presented with a graph indicating their current grade in the course compared to the goal they had set. It was important that this be visible to the students each and every time they logged into the course website so that they were frequently reminded of their progress.

Students could also view their progress in the class and on the current assignment in comparison to their peers. As mentioned in the literature (Section 2.1), people may be more likely to compare themselves to others who are similar to them. To address this, three groups of students were presented for comparison:

1. *Average students*, which was simply the average for the class as a whole, was presented so that students would be aware of how they compared to the class as a whole.
2. *Similar students*, which was the average of students with a goal set within 5% of that of the individual student, was included to provide students with a subsection of the class who were comparable in terms of their goals.

3. *Top students*, which was defined as the average of the students within the top 10% of the class, was incorporated with the hope of creating more competition between those who were a part of the group, and more motivation for other students to improve their grades in order to become part of the group.

### 3.3.2 Preliminary Design

Early on, it was decided that the best way to execute the proposed approach would be through the creation of a Moodle plug-in. The primary reasons for using Moodle were that it is open-source and familiar to the students who would be participating in the study. Therefore, development for the plug-in started with the capabilities and context of Moodle in mind.

Moodle allows for many types of plug-ins, each intended for different kinds of content (the full list of which is available at [http://docs.moodle.org/dev/Plugins](http://docs.moodle.org/dev/Plugins)). In order to narrow down the content which would be included in the plug-in, a list was created of all the pieces of information that could potentially be used in order to motivate students in the course. This list included the following items:

**To assist with achieving goals:**

- Reminders of upcoming help sessions
- Customized phrases of positive encouragement
- Goal for course set by the student
• Sliders displaying the students’ current grades relative to their goal grade

**To encourage individual progress:**
• Alerts for assignment due dates
• Completion status on each of the graded items in the course
• Graphs displaying the students’ progress on the current assignment

**To introduce competition and comparison with peers:**
• Graphs showing the proportion of students who were stuck on the current assignment
• A ranking of where students stood compared to their classmates
• Comparisons between the students’ progress and that of their classmates

Moodle’s *block* and *activity module* plug-in types were deemed most suitable for this application because they allow for almost any type of content and formatting. *Blocks* reside in the left-hand side panel of Moodle, and are small, collapsible, and positionable. They are limited in their capabilities and what they can display, but are easy to position wherever the instructor desires. *Activity modules* tend to take up more space, usually filling a page dedicated to them, and are much less limited in their capabilities. An example of block and module layout can be seen in figure 3.2

Several wireframes were drawn to explore how to lay out the content using both blocks and modules (see figures 3.3 and 3.4). Blocks alone simply did not provide enough space to present some of the information in a coherent and simple way. However, using only a module would mean that students would not be presented with the visualizations when they first log into the course Moodle page.
Figure 3.2: Example layout of Moodle blocks and modules.

Figure 3.3: Wireframes for brainstorming content and layout of Moodle block.

It was important that the students were able to see at least a portion of the information from the plug-in when they first entered the course website. This was essential in the event some students did not click on the link to the plug-in, they would still see some of the plug-in information. Seeing the plug-in with each login would also serve as a reminder that it was available to the students as well.
Figure 3.4: Wireframes for brainstorming content and layout of Moodle module.
It was the intention that the students would notice the information each time they logged in. Therefore, to give the plug-in presence on the course home page as well as having it contain enough information, both a Moodle block and a Moodle module were used. The content would be spread between the two, ideally with the most important information shown on the main course page through the block. The rest of the information would be contained in the module that would be a secondary page dedicated solely to plug-in content. Additional wireframes were created to explore the distribution of content between the block and module. One of these wireframes can be seen in 3.5.

Figure 3.5: Brainstorming distribution of content between block and module.

It was decided that the plug-in would allow students to set goals for their final grade in the class, present the students with their individual progress on the current assignment, and show them their progress in the course overall with respect to the goals they had set. It would also contain graphs that would allow students to compare their overall progress and assignment progress to that of the class.
These features were chosen because they seemed to have the best balance for simplicity of implementation, relevance, and impact on the students. The simple setting of goals and progress monitoring have shown to be motivational (Schunk, 1984). They also aid in the ability for self-assessment which has been deemed a necessary skill for learning (Mitrovic & Martin, 2006). The graphs displaying the progress of each student relative to the class were present to introduce an aspect of competition and comparison, which is theorized to be a major motivator (Festinger, 1954). It was hoped that students would be motivated to better themselves or help others because of this awareness of where they stood in contrast to their peers.

Another upside to these choices for visualization is that the majority of the data required was already present in the basic installation of Moodle. Therefore the plug-in could be easily replicated elsewhere without the need for additional data collection or programming of student models.

3.3.3 User Interface

The last design stages consisted of adjusting how the information would be presented. Usability, ease of understanding, and aesthetics were all considered. Because the students’ individual progress is likely more important to them than that of the class, the individual progress content was assigned to the block section of the plug-in. They could see how they were progressing towards their goal, but could only set their goal by going to the main module of the plug-in. This was done to give students incentive to visit the main plug-in page at least once. The comparison graphs, which would show a students their progress relative to the rest of the class, were located in the module, as these graphs required more space and background calculations. Thus, the final interface
for the plug-in contained the following features, split between the block and module:

In the module:

• Students could set a goal grade for themselves in the course. They could change their goal at any time and any number of times.
• Students could view the assignment progress of their classmates (as subsets of the class, not individually), along with a reminder of when the assignment was due.
• Students could compare their progress in the course to that of their classmates (again as subsets of the class).

In the block:

• Students could view their percentage of completion of the current assignment.
• Students were reminded of their grade relative to the goal they set for the course using another small graph with a short message.

Rather than the plug-in displaying progress data of the class as a whole, the researchers wanted to be able to provide the students with a comparison to smaller groups of students. In order to avoid the identification of individual students, but also to provide the student with a closer comparison to certain portions of the class, three subsets of the class were created as mentioned in Section 3.3.1. It was the hope of the researchers that having a more focused group for student comparison would create more competition between students and thus, a higher increase in motivation.

The module was split visually into two sections for organizational purposes. One section was allocated to class progress on the current assignment and the other for class progress in the course overall. The assignment section contained
three separate graphs, one for the progress of each subset of the class. It also displayed a reminder of how much time remained before the assignment was due.

The course progress section showed a sliding scale graph, indicating the individual student’s current grade, and showing markers of the grades for each of the class subsets. The use of the slider for the individual student was intended to highlight their progress amongst that of the class subsets. This area was also where the students were able to create and edit their goal for their grade in the class. The goal setting feature was purposely placed on this secondary page in order to give students a reason to click on the plug-in and therefore view these other visualizations. A screenshot of the final module design can be seen in figure 3.6. (The block and module are displayed separately only to allow room for readability. Students were able to see both at once in the plug-in.)

The block used another circular graph to display assignment progress and a small horizontal graph to display the student’s current grade in the course with a marker showing the goal they had set. They were also provided with a message stating the exact difference between their grade and their goal. A screenshot of the final block design can be seen in figure 3.7.

The graphs used in the Progress Tracker were chosen for their balance of aesthetics and intuitiveness. The circular assignment progress graphs resemble a loading bar, which normally represent a percentage towards completion, where at the final stage it is expected to reach 100%. It was assumed that students aim to complete their assignments, and therefore this form of visualization seemed suitable. Overall progress was shown as a single horizontal bar graph with sliding markers to mark important levels. The individual progress bar contained a sliding marker positioned at the grade the students had set for their goal. The overall class progress used a similar graph where the “fullness”
Figure 3.6: Screenshot of what the students saw when they accessed the progress tracker.

of the bar indicated the individual student’s progress while sliding markers indicated the progress of each of the class subsets.

The graphic used for the circular graphs was designed by Piotr Kwiatkowski from a blog called Better2Web (Kwiatkowski & Jeffery, 2012) and was provided free of charge for use in personal and commercial projects. The horizontal graphs were designed and created by the researcher to match the look of the graphs by Kwiatkowski.

Several help bubbles were added to explain the subsets and graphs to the students using the progress tracker. The help bubbles were located anywhere there was a grey question mark in a circle. If students hovered their mouse over
one of the icons, an explanation of that element would appear. An example of one of the help bubbles can be seen in figure 3.8.

Basic user testing was done on the Progress Tracker by fellow graduate students. Their comments and concerns led to some minor changes and additions. The help bubbles (as described above) were added, mostly to clarify the meaning of the class subset labels. Error checking was added to prevent unexpected input for goals and error messages were added so that the user knew why their graphs were not visible in some situations (such as when there were less than five similar students to compare to).
3.3.4 Implementation

The Moodle web application is coded using PHP, therefore the Progress Tracker plug-in was implemented in PHP as well, with Javascript used to animate the graphs. The existing Moodle PostgreSQL database was used to store information, and the code was structured using Moodle’s activity module and block development guides available at http://docs.moodle.org/dev/Activity_modules and http://docs.moodle.org/dev/Blocks.
3.3.4.i Setting Goals

It was not possible for students to change their goals directly from the block portion of the plug-in due to Moodle’s blocks being more limited in their capabilities. As such, students were able to set their goals from the module. There was no limit to how many times students could change their goal.

Each time students set or changed their goal, a new record was added to a table called “progresstracker_goals”. These records contained the ID of the course the goal was set for, the user ID of the student who set it, the goal itself, and a timestamp of when it was set or changed.

In order to avoid the entire page refreshing each time a goal was set, an AJAX POST was used to handle communication between Javascript on the

![Diagram of communication process for successful goal change.]

Figure 3.9: Communication process for successful goal change.

In order to avoid the entire page refreshing each time a goal was set, an AJAX POST was used to handle communication between Javascript on the
client side and PHP on the server side. Figure 3.9 illustrates the communication executed during a successful goal change. First the user enters a goal, then clicks the “Change” button. The button click action is caught by Javascript on the client side. The Javascript process then creates a JSON object containing the student’s goal, user ID and the course ID (which were initialized when the plug-in was launched). This JSON object is sent using AJAX’s POST method over the server side PHP. The PHP process saves the new goal record to the database, creates a feedback message for the user, calculates the new grade and assignment progress for “similar” students, then re-encodes all of this in a new JSON object which is echoed back to the client side. Finally, the client side Javascript displays a message telling the user that the goal was successfully changed, then uses JQuery to update the graphs according to the new “similar” student statistics.

3.3.4.ii Calculation of Class Subsets

As mentioned in Section 3.3.1, the class was divided into three subsets (average, similar, and top students) to allow for closer peer comparison. The students belonging to these subsets were found according to the following three algorithms (1, 2, and 3.) In each case, if the subset contained less than five students then the graph would not display any results. This was done to respect the students’ privacy and help avoid results that could be skewed if based on only one student. (For instance, if the class size was small, the top students progress graph could show the grade of only the single highest graded student, thus giving away that student’s grade to anyone who knew who was doing best in the class.) Five was chosen for the threshold because the researcher felt it was a large enough group to compare to without having the results too biased by one student, and small enough that the plug-in could still function in smaller class sizes.
Algorithm 1: Select average students in the class.

Require: courseid

avgStudents ← SELECT students WHERE course = courseid

return avgStudents

Algorithm 2: Select similar students in the class.

Require: courseid, goal

span ← 5

simStudents ← SELECT students WHERE course = courseid
AND goal BETWEEN goal + span AND goal − span

while simStudents < 5 do
    span ← span + 1
    simStudents ← SELECT students WHERE course = courseid
    AND goal BETWEEN goal + span AND goal − span
    if span > 25 then return 0
    end if
end while

return simStudents

The subset of average students contained all students in the class, and was therefore found using single select statement on several joined tables in the database. A simplified version of this database call is shown in algorithm 1.

The subset of students similar to the current student was based on the current student’s goal. Initially the subset contained all students in the class that had a goal within 5% of the current student. However, if this resulted in a subset of less than 5 students, the goal range was widened until the subset contained more than five students (with a maximum range of 25% in either direction). Algorithm 2 shows this process.

Lastly, the subset of the top students in the class consisted of those with grades in the top 10%. Algorithm 3 shows this process.
Algorithm 3: Select top students in the class.

Require: courseid
limit ← total students in course / 10
if limit < 5 then limit ← 5
end if

topStudents ← SELECT students WHERE course = courseid ORDERED BY grade
LIMIT at limit
return topStudents

3.3.4.iii Assignment Progress and Overall Progress Graphs

Data from an existing Moodle plug-in called “Checklist” was used to determine the progress of each student on the assignment. The Checklist plug-in allows an instructor to add a set of items to a “to-do list” that the students can view and complete at their own pace. In CIS2500, the instructor used this plug-in to create checklists for each of the assignments. This helped to provide the students with some additional guidance about what they needed to do in order to complete the assignment as well as provide a tool to aid with time management. The class had been using these checklists from the start of the course, and was therefore familiar with them by the third assignment when the Progress Tracker was added.

Items in the assignment checklist were variable in the expected time commitment to the assignment overall. For instance one item could be simply reading the assignment specification whereas another could be coding a large portion of the assignment. In order to accurately represent the completion percentage of the assignment, each checklist item was assigned an expected time commitment. This time commitment was used when calculating how much of the assignment a student had completed. One example of a checklist item with this format could be “Add error checking and handling to your code (10% of
Algorithm 4: Calculation of assignment progress for class subsets.

Require: students, checklist
for all items in checklist do
    timeCommitment ← percentage specified in item title
    completed ← COUNT students WHERE item is completed
    avgProgress ← avgProgress + ((timeCommitment × completed)/students)
end for
return avgProgress

Algorithm 5: Calculation of grade for class subsets.

Require: students
for all students do
    avgGrade ← avgGrade + grade of student
end for
avgGrade ← avgGrade/students
return avgGrade

your time).” So if this item were checked, it would increase a student’s assignment progress graph by 10%. If all items were checked, the assignment progress would be 100%.

The assignment progress was calculated as described above for each student. The progress of the logged in student as an individual was displayed in the block portion of the plug-in. The class assignment progress graphs displayed the average assignment progress for each of the three subsets as described in 3.3.3. The progress for these comparison graphs was calculated according to Algorithm 4 for each of the three subsets.

To display the overall progress in the class (the student’s current grade), each student’s grades were selected from the Moodle database. The block portion of the Progress Tracker displayed the individual student’s grade, while the module showed that grade along with that of each of the class subsets along a single horizontal progress bar as described in 3.3.3. The grades for the three class subsets were calculated as shown in Algorithm 5.
3.4 Experimental Design

The goal for the experiment was to see whether the addition of the progress tracker plug-in had an effect, either positive or negative, on the students’ engagement in the class. The experiment also tried to identify the flaws and strengths of the Progress Tracker, and consequently, ways in which it could be improved so that it would be more beneficial to the students.

To acquire this information two questions were put forward for examination through analysis of the data:

• Question 1: Can students’ engagement in an e-learning environment be improved by providing comparative visualizations of the students’ progress?

• Question 2: Do students’ believe that the visualization helped improve their motivation in the e-learning environment?

As mentioned, the plug-in was implemented for the third and final assignment in CIS*2500. It was incorporated at this point in the semester so that the data for the first two assignments could be compared to that of the third assignment during which the Progress Tracker was available. Therefore Question 1 was answered by comparing the results of attributes known to indicate engagement across each of the three assignments. More favourable values in these attributes were assumed (based on the studies discussed in Section 2.4) to point to an increase in engagement, thus the results during assignments one and two (before the plug-in was deployed) were compared to those during assignment three (when the plug-in was available). These same attributes were also compared between the participants who had used the plug-in for goal-setting against those who did not.
Question 2 was addressed primarily through the contents of the Student Course Engagement Questionnaire (SCEQ) (Handelsman et al., 2005) and Progress Tracker Feedback Questionnaire (PTFQ) that were administered to students during the last three weeks of the semester. The questionnaires were available to all students in the class regardless of their participation in the study. All questions were posed as positive statements, and were asked on a 5-point agree-disagree scale. The questions in the SCEQ were not altered at all.

The PTFQ was created by the researcher to collect the students’ opinions of the Progress Tracker. It contained 10 agree-disagree questions in the same format as the SCEQ and 4 open-ended questions to allow students to provide comments. This survey data identified the aspects of the plug-in that the students felt were useful as well as the areas that they thought needed improvement. It also provided some insight into whether students believed the plug-in instigated any motivation driving factors, such as competition or comparison. The rationale for each question asked in the PTFQ can be seen in Table A.3 in Appendix A.

In addition, further exploration of the data was done to attempt to uncover further information that could help better induce engagement through information visualization. The use of clustering (Section 4.5) as well as observing trends (Section 4.3) in the data helped to identify student behaviours which led or may lead to success. These behaviours can then be encouraged in future designs of visualization plug-ins.
3.5 Experimental Procedure

3.5.1 Plug-in Deployment

The Progress Tracker was added to the CIS*2500 Moodle site on March 8th, 2013 (week 8 of the semester), at the same time as the third assignment for the class was posted. The students in CIS*2500 were still relatively new to programming assignments, and the assignments in this class were significantly more time consuming than those in the class preceding it. For this reason, there was speculation that the data for the first assignment might be skewed because the students were simply not expecting to spend so much time to work on it, and therefore started later than they would have otherwise.

3.5.2 Data Collection

Data collection commenced shortly after the students’ official grades had been released by the university. This ensured that there could be no bias in the grading of the class based on which students had chosen to participate or not participate in the study. After the grades were released, the consent forms for the students who wished to participate in the study were collected and a list of participants was compiled.

The data was collected from five sources as discussed in Section 3.2.4. After the completion of the semester a SQL dump of the Moodle database was done to preserve the data collected throughout the course of CIS*2500. The CIS*2500 course was also kept on the Moodle website in a “Past Courses” section so that data could be gathered directly from the site as well.

The participants’ goal setting information was obtained from this SQL dump through database queries on the tables used for the Progress Tracker plug-in. This data included every goal (or change in goal) set by students in
the class, along with a corresponding timestamp of when that goal was entered and the student’s name.

Log files containing each participant’s interactions with Moodle were downloaded from Moodle’s log reporting feature (located under Course administration >Reports >Logs.) The log data included information on the actions of each participant in Moodle throughout the course of the semester. There was too much log data for Moodle to create a single log file for the class (it would hang when attempted), therefore the data was downloaded as an individual comma separated value file for each participant. While these activity logs also could have been retrieved from the SQL dump of the database, downloading them using Moodle’s log report feature allowed them to be in a more human-readable format right from the start, saving some data pre-processing time.

The participants’ responses to the SCEQ and PTFQ were also downloaded from the Moodle website. The questionnaires were created using Moodle’s built-in feedback activity module which includes an analysis of the results downloadable in an Excel spreadsheet format.

Grades for each component of the class as well as the final grades were provided by the instructor of CIS*2500. Responses to questions asked in class (both clickers and written), and Git logs were also acquired from the instructor. These indicated the students’ attendance to lectures.

3.5.3 Participation

Of the 192 students enrolled in CIS*2500 at the time of recruitment, 75 volunteered to be participants in the study. The level of participation was good, considering that recruitment was done only during the final week of classes. However, the resulting sample of students was not entirely representative of the class as a whole. The grade average of the participants was slightly higher.
at 75.96% whereas the class average was 71.26%. This could be due to the time at which recruitment was carried out. Towards the end of the semester, less students attend lectures, and those that are still attending tend to be those that are more engaged and still believe that they can pass the course (Van Blerkom, 2001). Therefore the students that were doing poorly or did not care about the class, may not have been a part of the audience during the recruitment sessions.

![Proportion of participating students](image)

**Figure 3.10:** Proportion of participating students and their contributions.

Of the 75 participants, 58 had opted to use the Progress Tracker to set a goal for themselves; 22 of the participants completed the SCEQ and PTFQ which provided their self-reported engagement level and feedback and satisfaction with the Progress Tracker; 18 of the participants both used the Progress Tracker and completed the questionnaires. Figure 3.10 shows a graphical representation of these proportions.
3.5.4 Data Cleaning and Amalgamation

A fair amount of data cleaning was needed in order to transform the raw data into a data set that met the ethics requirements outlined in the Data Collection Plan (table B.1 in Appendix B) and that would be suitable for analysis with R and Weka. This included transferring percentage grades to letter grades (see Table A.2), replacing identifying information with random participant IDs, rounding timestamps to the nearest hour, counting Moodle accesses, extracting data from log files and so on. Each portion of data mentioned in Section 3.5.2 was aggregated into one main data set using a combination of scripting, spreadsheet calculations, and manual entry. In the end, this main data set contained one row per participant and a total of 86 attributes, or columns (including the randomly generated participant ID).

The access log data downloaded from Moodle was initially in the format seen in figure 3.11, with one log file per participant. Since this format was not very conducive to analysis, two scripts were written in order to transform the log data into something more useful.

The first script merged all of the logs into a single file then counted the total number of actions of each type (course, forum, checklist, quiz, assignment, lesson, progress tracker, and feedback) for every day of the semester. This resulted in a square data set (one row per day and with no missing values), kept separate from the main set, and provided information about how the class as a whole was accessing Moodle.

The second script counted the different types of actions within specified time spans for each of the participants (checklist updates for each assignment, overall access during each assignment, forum views and so on). Assignment start times were recorded using the script as the earliest “assignment view”
event for each assignment. Assignment submission times were recorded as the last “assignment submit” event for each assignment. These variables calculated via the script for each participant were included as part of the main data set. Note that these assignment submissions did not include the actual files that made up each student’s project. The submissions made using Moodle were only “collaboration statements” (a paragraph stating the student had not participated in any academic misconduct) that students were required to submit along with their projects and will be referred to as “Moodle submits”. The actual files were submitted through a Git, referred to as “Git submits”, and are explained below.

Git logs, which also had one log file per participant, were merged into a single file using another bash script to make it easier to extract the desired information for all of the participants at once. The primary data collected from the Git logs was commit information. Each time a student does a “commit” a snapshot is taken of the files they are submitting and is recorded in their project history. The Git logs record every commit from every student, providing information about when and how often they made changes to their project as well as what those changes were. The total number of commits for each assignment, as well as the date for the final commit were found for each participant and included in the main data set.

A similar process was followed for the goal logs recorded by the Progress Tracker plug-in. Each participant’s first goal was recorded as the first goal that remained stable for at least five minutes. This was done since several students appeared to rapidly change their goals just to see how the graphs responded as a result. The last goal, and total number of goal changes (at least 5 minutes apart) were also recorded for each of the participants.
Figure 3.11: Example contents of Moodle Logs for a participant.

<table>
<thead>
<tr>
<th>Course</th>
<th>Time</th>
<th>IP address</th>
<th>User full name</th>
<th>Action Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIS500</td>
<td>16 March 2013, 11:02 AM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>“Course view (<a href="http://bucky.socs.uoguelph.ca/course/view.php?id=2">http://bucky.socs.uoguelph.ca/course/view.php?id=2</a>), &quot;CIS 5000 Intermediate Programming&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>11 March 2013, 11:18 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Course view (<a href="http://bucky.socs.uoguelph.ca/course/view.php?id=2">http://bucky.socs.uoguelph.ca/course/view.php?id=2</a>), &quot;CIS 5000 Intermediate Programming&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>11 March 2013, 5:17 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Course view (<a href="http://bucky.socs.uoguelph.ca/course/view.php?id=2">http://bucky.socs.uoguelph.ca/course/view.php?id=2</a>), &quot;CIS 5000 Intermediate Programming&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>13 March 2013, 5:17 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Quiz view (<a href="http://bucky.socs.uoguelph.ca/mod/quiz/view.php?id=45">http://bucky.socs.uoguelph.ca/mod/quiz/view.php?id=45</a>), &quot;Lab 7: Linked Lists&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>13 March 2013, 5:17 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Course view (<a href="http://bucky.socs.uoguelph.ca/course/view.php?id=2">http://bucky.socs.uoguelph.ca/course/view.php?id=2</a>), &quot;CIS 5000 Intermediate Programming&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>13 March 2013, 5:17 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Page view (<a href="http://bucky.socs.uoguelph.ca/mod/page/view.php?id=45">http://bucky.socs.uoguelph.ca/mod/page/view.php?id=45</a>), Tutorials&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>2 March 2013, 4:04 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Progress tracker view (<a href="https://bucky.socs.uoguelph.ca/mod/course/tools.php?id=27">https://bucky.socs.uoguelph.ca/mod/course/tools.php?id=27</a>), 1&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>2 March 2013, 4:04 PM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Course view (<a href="http://bucky.socs.uoguelph.ca/course/view.php?id=2">http://bucky.socs.uoguelph.ca/course/view.php?id=2</a>), &quot;CIS 5000 Intermediate Programming&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>2 March 2013, 2:45 AM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Forum view discussion (<a href="http://bucky.socs.uoguelph.ca/mod/forum/discuss.php?id=381">http://bucky.socs.uoguelph.ca/mod/forum/discuss.php?id=381</a>), &quot;Lab 6 write&quot;</td>
</tr>
<tr>
<td>CIS500</td>
<td>2 March 2013, 2:45 AM</td>
<td>131.184.1.1</td>
<td>John Doe</td>
<td>&quot;Forum view forum (<a href="http://bucky.socs.uoguelph.ca/mod/forum/view.php?id=7">http://bucky.socs.uoguelph.ca/mod/forum/view.php?id=7</a>), &quot;General Discussions&quot;</td>
</tr>
</tbody>
</table>
Attendance information was collected from the clicker (and written answer) grades from each lecture. If a student had any grade at all for that lecture (meaning they attempted to answer at least one question), then they were counted as having attended the lecture.

Scores from the SCEQ were calculated as percentages. Each question had a score range from 0 to 4, so with a total of 26 questions, this gave a total score between 0 and 104. The scores for each of the 4 measures within the questionnaire (skills, emotional, participation, and performance) were recorded as well. The final three questions from the survey, which were scored differently, were included as separate attributes.

Once the data had been merged into a single data set it was then modified to comply with the ethics requirements for this experiment. In all cases, identifying information was removed from the data and was replaced by participant ID numbers. The information removed included names, student IDs (the first part of a student’s University email address), student numbers, IP addresses, and email addresses. Additional precautions were taken to reduce the likelihood of being able to backtrace the data and link a participant to a student. These precautions included rounding any timestamps to the nearest hour and converting all grades to a letter grade scale (including SCEQ scores). A detailed description of the alterations to each section of the data can be seen in Appendix B Table B.1.

In addition to removing the students’ personal information, the data was also scrubbed for completeness and accuracy by manually examining each entry. Incomplete or clearly inaccurate (such as an assignment submission date before the assignment had been posted) records were either corrected manually or discarded. For example, one cause of missing information that was encountered was when an assignment was submitted directly to the instructor rather
than through the Moodle website, thus resulting in a participant having no submission time, yet still having received a grade for the assignment. If the missing information was available from the instructor it was corrected manually in the data. Missing data was considered valid in most cases (such as not having set a goal or not having a Git submission time) and was therefore not replaced.

Some additional attributes were added to the data after it had been cleaned, mainly as summary attributes. Averages for assignments, quizzes, and lab grades were calculated for the study using the same weighting as in the course (therefore maintaining the importance of each towards the final grade). An attribute was added for whether participants had met their goal or not and a total attendance column was added containing the total number of recorded lectures attended by each participant. Table 3.1 contains descriptions of a selection of attributes. The final data set, with descriptions of all 86 attributes and details of how they were derived (if applicable), can be seen in Appendix C, Table C.1.

To allow for numerical analysis, a second version of the main data set was created using numerical percentage grades rather than letter grades and binary values in place of yes/no attributes. The grades were converted to the median number in the range of percentage grades corresponding with each letter grade (see Table A.2 in Appendix A). For example, an A+ corresponds to percentage grades from 90% to 100%, therefore every A+ was changed to a grade of 95% for the numerical data set.

3.5.5 Data Analysis

The R statistical analysis environment (R Core Team, 2014) was used to examine this numeric data set. The portion of analysis done with R included the
<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>Description</th>
<th>Post-Cleaning Derivation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>First goal</td>
<td>The first steady goal set by the participant using the progress tracker. The goal was recorded if it remained steady for 5 minutes. This was done because some participants tended to rapidly alter their goal, likely in order to view the resulting changes in the graphs.</td>
<td>Converted to letter grade</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Skills</td>
<td>The participant’s score on questions 1 to 9 of the SCEQ.</td>
<td>skillsScore/36</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Emotional</td>
<td>The participant’s score on questions 10 to 14 of the SCEQ.</td>
<td>emotionalScore/20</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Participation</td>
<td>The participant’s score on questions 15 to 21 of the SCEQ.</td>
<td>participationScore/28</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Performance</td>
<td>The participant’s score on questions 22 to 24 of the SCEQ.</td>
<td>performanceScore/12</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Final</td>
<td>The participant’s final grade in the class.</td>
<td>Original data</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Assign avg</td>
<td>The average of the participant’s three assignment grades.</td>
<td>(A1 + A2 + A3)/3</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Lab avg</td>
<td>The weighted average of the participant’s three lab exam grades (worth 5%, 10%, and 10%).</td>
<td>(LE1 × 0.5 + LE2 + LE3)/2.5</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Quiz avg</td>
<td>The weighted average of the participant’s two quiz grades (worth 5% and 10%).</td>
<td>(Q1 × 0.5 + Q2)/1.5</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Clicker</td>
<td>The participant’s participation grade in the class, contributed to by their “clicker” activity during lectures.</td>
<td>Original data</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Exam</td>
<td>The participant’s grade on the final exam.</td>
<td>Original data</td>
<td>A+ to F</td>
</tr>
<tr>
<td>PT views</td>
<td>The total number of times the participant viewed the main progress tracker page throughout the observation period.</td>
<td>Original data</td>
<td>0-34</td>
</tr>
<tr>
<td>Forum views</td>
<td>The total number of times the participant viewed the forums throughout the observation period.</td>
<td>Original data</td>
<td>34-3903</td>
</tr>
<tr>
<td>Total access</td>
<td>The total number of times the participant accessed the CIS*2500 Moodle page.</td>
<td>Original data</td>
<td>310 to 7385</td>
</tr>
<tr>
<td>Total attended</td>
<td>The total number of lectures attended by the participant (of those with attendance information).</td>
<td>( \sum L1.1 + L1.2 + \ldots + L12.3 )</td>
<td>2 to 25</td>
</tr>
</tbody>
</table>

Table 3.1: Sample of attribute descriptions
calculation of correlations within the data, statistics of significance, analysis of variance, and tests for significant differences across multiple test attempts (such as access during assignments one, two, and three). R was also used to create graphs and charts for illustration of the data.

To further explore the data, the Weka data mining software (Hall et al., 2009) was utilized. Weka allows the user to apply a wide range of pre-programmed classification and clustering algorithms to a set of data, making it easy to investigate predictive hypothesis in the data. In this case, Weka was used to create clusters within the data so that the researchers could better understand what variables influence each other when looking at students in an e-learning setting. Clusters were made using both the numerical and letter-grade data sets, as they produced different results. Weka does not interpret an A+ as being similar to an A, however it knows that a grade of 95% is similar to a 87%. A description of the clustering algorithm used and the parameters chosen can be seen in Section 3.2.4.
Chapter 4

Results and Discussion

Log data from the Moodle LMS was used in combination with Git logs, attendance data, grades, and self-reported measures of engagement and satisfaction to determine if providing students with visualizations of their progress in the course along with that of their classmates improves engagement within the e-learning system. The data was collected from 75 voluntary participants in a second year computer science class. This chapter presents the results of data mining and analysis of the collected information.

4.1 Summary Statistics of the Data Collected

Basic summary statistics were calculated for each attribute in the collected data. Table 4.1 shows the number of measurements, minimum, median, maximum, mean, and standard deviation of the same sample of attributes presented in table 3.1 in the previous chapter. The full table of summary statistics for all attributes can be seen in table C.2 in Appendix C.

Access data from the Moodle LMS was recorded throughout the span of the semester. Recorded accesses were categorized according to the components of
Table 4.1: Data collected

Moodle used in CIS*2500. These components included course content, forums, checklists, quizzes, assignments, and lessons. This data provides information about which aspects of the e-learning system students interacted with in response to course events and when they accessed those different features.

Figure 4.1 illustrates the average number of each type of Moodle access (course, forum, checklist, quiz, assignment, or lesson view) on each day of the semester. The due dates of the assignments, the day of the exam, and the day that the Progress Tracker was deployed are marked on the graph with a vertical line on the day of the event.

The graph shows the participants’ use of Moodle components in response to the course events indicated. Before each assignment large spikes in forum access are seen, with the size of the spike increasing with each subsequent assignment. Increases in course content, assignment, and checklist access are
Figure 4.1: Participants’ Moodle activity for each day of CIS*2500
Figure 4.2: Participants’ Moodle activity for each week of CIS*2500
also seen around the due dates of each assignment. Lesson access increased with assignments two and three, but no increase was seen for assignment one. Finally, spikes in quiz, course, forum, and lesson access are seen just before the final exam.

Figure 4.2 displays the total access separated by type during each week of the semester. As with Figure 4.1 the weeks during which assignments were due and the Progress Tracker was deployed are indicated in the weeks shown with a grey gradient. Here the trends in type of access are more easily recognized. High spikes in course and forum access are seen before the conclusion of the final two assignments. The waves of checklist access also seem to correspond with the timing of the assignments.

4.2 Survey Results

The SCEQ and PTFQ were posted at the start of the last week of classes. The students in CIS*2500 had the remaining three weeks of the semester to complete the questionnaires. The SCEQ gathers self-reported engagement measures in four areas: skills, emotional, participation, and performance. For data analysis, the results of the survey were converted to percentage scores. To quantify the results, each question was scored from 0 to 4, with a score of 0 corresponding to “strongly-disagree” and 4 corresponding to “strongly-agree”. This provided a final SCEQ score between 0 and 104, which was then converted to a percentage grade for continuity with the other graded items in the data set. Scores were also recorded for each of the four individual areas. The skills measure was reflected by questions 1 to 9, emotional was reflected by questions 10 to 14, participation was reflected by questions 15 to 21, and performance was reflected by questions 22 to 24.
Table 4.2: SCEQ questions and results.

Table 4.2 contains the results gathered from the SCEQ. Responses to both “strongly-agree” and “agree” were added to form the numbers in the Agree column. Likewise, responses from “strongly-disagree” and “disagree” were added to form the Disagree column. The statistics for neutral answers can be seen in Figure 4.3. The SCEQ is divided into four sub-areas of engagement, “skills”, “emotional”, “participation”, and “performance”. The questions in the table have been partitioned into the sub-area that their scores contribute to.
Figure 4.3: Results from the Likert portion of the SCEQ survey

Figure 4.3 shows a graphical illustration of the responses to the SCEQ. The figures show each question asked in the questionnaires and the corresponding responses to that question. The orange and yellow end of the bars indicate the percentage of “strongly disagree” and “disagree” responses, and the turquoise and lime end of the bars indicate the percentage of “strongly agree” and “agree”...
responses. The neutral colour, on which the bars are centred show the percentage of “neutral” responses. Exact percentages for disagree, neutral, and agree responses are marked to the left, center, and right of the bars respectively.

The results from the SCEQ indicate that almost all the participants felt that they regularly put forth effort, yet only 50% said that they participated in discussions. Only 32% of participants said that they raise their hand in class and only 32% said that they go to office hours.

Correlations between the SCEQ scores and other data recorded during the course can be seen in Section 4.3, Figure C.2 in Appendix C.

Table 4.3: PTFQ questions and results.

<table>
<thead>
<tr>
<th>PTFQ Question</th>
<th>Agree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) I regularly checked the graphs in the Progress Tracker</td>
<td>68.4%</td>
<td>10.5%</td>
</tr>
<tr>
<td>2) I liked being able to see my progress towards my goal</td>
<td>84.2%</td>
<td>0</td>
</tr>
<tr>
<td>3) I liked being able to compare my progress on the assignment to the rest of</td>
<td>68.4%</td>
<td>5.3%</td>
</tr>
<tr>
<td>the class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) I liked being able to compare my overall grade to the rest of the class</td>
<td>89.5%</td>
<td>0</td>
</tr>
<tr>
<td>5) The information provided through the plug-in made me more competitive</td>
<td>36.8%</td>
<td>21.1%</td>
</tr>
<tr>
<td>6) The information provided through the plug-in made me more engaged in the</td>
<td>52.6%</td>
<td>10.5%</td>
</tr>
<tr>
<td>assignment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) Knowing how the rest of the class was doing helped motivate me</td>
<td>57.8%</td>
<td>10.5%</td>
</tr>
<tr>
<td>8) Being reminded of my own progress towards my goal helped motivate me</td>
<td>78.9%</td>
<td>5.2%</td>
</tr>
<tr>
<td>9) I liked the look/aesthetics of the plug-in</td>
<td>78.9%</td>
<td>5.3%</td>
</tr>
<tr>
<td>10) I would use this plug-in in all my courses</td>
<td>78.9%</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.3 contains the responses given to the PTFQ that was included along with the SCEQ. Figure 4.4 shows the graphical illustration of these results (which are to be interpreted the same way as Figure 4.3). The feedback for the Progress Tracker was very positive. As the figure shows, the majority of responses were at the agree end of the scale for all questions asked. Only one of the participants who completed the survey reported not liking that they could compare their progress to the rest of the class (unfortunately they did
not specify why in the open-ended questions the end) and only one disagreed to the statement “being reminded of their progress helped to motivate me”. The students were also asked open-ended questions about what they would change in the plug-in and what aspects they did not like. These results give some better ideas for future improvements on the Progress Tracker and will be discussed in Section 5.1.

![Figure 4.4: Results from the likert portion of the Progress Tracker survey](image)

### 4.3 Data Exploration

The data were first observed through pairwise scatter plots of the variables. The scatter plots appeared to show several linear relationships between the variables but no obvious non-linear relationships. The three scatter plots in Figure 4.5 show examples of the main trends seen in the data. To analyze the linear correlations, the data were explored using Spearman’s rank-order correlation. Spearman’s correlation was chosen since the data did not follow a normal distribution. Much of the data consisted of counts, which are discrete,
and therefore not normally distributed. Graded items could follow a somewhat normal distribution, but in this case observations of the histograms of these variables showed that they did not (refer to Appendix C.1 for histograms of the graded item attributes). R’s “pairwise complete observations” option was selected so that missing data would be dropped only between the two variables being compared rather than dropping that participant’s data entirely.

(a) Final goal versus final grade.  (b) PT views versus final grade.  (c) PT views versus A3 commits.

![Scatter plots](image)

Figure 4.5: Examples of scatter plots.

Binary and categorical attributes were not included in the correlation matrix as Spearman rank-order correlation cannot be used to calculate correlations between these data types. Figure 4.6 shows a summary of the resulting correlations. The full set of correlations between all ordinal and count attributes in the data set can be seen in Figure C.2 of Appendix C. In these figures, each attribute is shown along both the X and Y axis. The correlation between any two attributes is displayed at the intersection of those attribute labels in the matrix. The bottom half displays the correlation, where the strength and direction of the correlation are indicated by the colour and direction of the ellipse. A positive correlation is shown by a right-slanting ellipse and a negative correlation is shown by a left-slanting ellipse. The strength of the correlation
is reflected by the intensity of the colour of the ellipse; darker aqua indicates a strong positive correlation whereas darker orange indicates a strong negative correlation. The top half of the figure shows Spearman’s rank correlation coefficient which ranges between -1 and 1, with -1 indicating a negative linear correlation, 0 indicating no linear correlation, and 1 indicating a positive linear correlation.

Figure 4.6: Spearman rank-order correlation for a portion of the aggregated data.
The resulting correlation matrix showed that the summary grade attributes ("assign.avg", "lab.avg", "quiz.avg", and "exam") were highly correlated between one another ($0.45 \leq r_s(75) \leq 0.53, p < 0.0001$). In the full correlation matrix (Figure C.2) it can be seen that all graded attributes were positively correlated except for the clicker and lab exam three (LE3) grades ($r_s(75) = -0.10, p > 0.05$). This however, makes sense because LE3 was based purely on the application of programming skill rather than class material unlike the other graded items.

All Moodle access attributes ("pt.views", "forum.views", and "total.access" to "grades.access" shown in Figure C.2) were closely correlated to one another ($0.28 \leq r_s(75) \leq 0.92, p < 0.016$). The "total.access" attribute is made up almost entirely of the other access attributes, and therefore had to be strongly correlated to them. The strong positive correlations between the access attributes were expected. It makes sense that if students access one aspect of the LMS then they are likely to access other aspects. Or if they access the LMS more during assignment one then others it would be expected that they would also have access more during the other assignments.

The attributes representing the SCEQ data ("skills", "emotional", "participation", and "performance") showed some very interesting and unexpected relationships. It was expected that higher SCEQ scores in any category would be positively correlated with grades and any attributes that are known to indicate engagement (access variables, submit times, start times, checklist updates, Git commits, and attendance). However, the correlation matrix indicated that the emotional and participation measures in the SCEQ show several strong negative correlations in these areas. For instance, emotional versus A1 access ($r_s(22) = -0.59, p < 0.005$) and participation versus PT views ($r_s(22) = -0.54, p < 0.01$). See Figure C.2 for more of these relationships.
Surprisingly, higher scores in the skills area (regular studying, effort, organization, listening, and attendance) indicated lower grades on LE3 ($r_s(22) = -0.44, p < 0.05$), and Q1 ($r_s(22) = -0.67, p < 0.001$). High skills scores also pointed to high classroom participation and more Moodle access, especially during assignments (however these results were not significant $0.22 \leq r_s(22) \leq 0.40, p > 0.07$). Although not statistically significant this could suggest the self-reports of participants' skills were accurate since attendance and study habits were among the skills being evaluated. Higher scores in the skills area also pointed to lower scores in the emotional and performance area of the SCEQ (again not significant $r_s(22) = -0.26, p > 0.2$, $r_s(22) = -0.36, p > 0.1$ respectively), which were both positively correlated with all graded items. This could suggest that students who are less confident (-performance) and who do not have strong interests in the material (emotional) study more to make up for their lack of pre-existing knowledge.

The emotional area of the SCEQ (making course material relevant to life, strong desire to learn the material, and thinking about the course between lectures) was positively correlated with A2 ($r_s(22) = 0.51, p < 0.02$) and had positive, but non-significant, correlations with all other graded items except LE1 and clicker grades. This relationship is consistent with the skills versus emotional correlation. A high emotional score seems to indicate students who are interested in the material, and therefore might already have the associated knowledge and skills. The negative correlation with clicker grades (earned in lectures) would likely indicate that these students did not feel they needed to attend a lot of lectures to do well. The emotional measure was negatively correlated with every attribute representing Moodle access, supporting that speculation ($-0.59 \leq r_s(22) \leq -0.15$). If these students believe that they
already know the material, then they may have less need to reference Moodle in order to complete their assignments.

The results from the participation area (raising hands in class, asking questions, participating in discussions, and attending office hours) showed some interesting correlations, however the majority were statistically non-significant. From the correlations shown in Figure C.2 it appears that higher participation scores indicate less LMS access and have no indication of increased lecture attendance. Therefore, students who report that they actively participate do not necessarily attend any more classroom lectures and are less likely to access the LMS. Since the majority of statements in the participation section relied on classroom attendance this measurement of the SCEQ seemed to be biased by skewed self-evaluations. High participation scores were correlated with low performance scores and low final grades. Perhaps students who did poorly felt embarrassed to indicate their actual participation levels, or perhaps they had skewed expectations of themselves.

The performance area of the SCEQ was correlated as expected. A high performance score was strongly correlated with high grades (performance vs final grade $r_s(22) = 0.76, p < 0.0001$). Performance was also positively correlated with lecture attendance ($r_s(22) = 0.43, p < 0.05$). Although non-significantly, it was positively correlated with the emotional area ($r_s(22) = 0.36, p > 0.1$) and negatively correlated with skills ($r_s(22) = -0.36, p > 0.1$). Therefore, high performance scores suggested students who were confident and interested in the course material, and thus were satisfied with how well they were doing in the course.

All three of these areas (and thus the overall SCEQ score) had a negative correlation to the number of Progress Tracker views, albeit a non-significant
one \( r_s(22) = -0.38, p > 0.08 \). Indicating that those who are more engaged according to the questionnaire are less likely to use the plug-in.

<table>
<thead>
<tr>
<th>Spearmann correlation</th>
<th>A1 vs. A1 access</th>
<th>A2 vs. A2 access</th>
<th>A3 vs. A3 access</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.02</td>
<td>0.07</td>
<td>0.23</td>
</tr>
<tr>
<td>P-value</td>
<td>0.8333</td>
<td>0.5535</td>
<td>0.0438</td>
</tr>
</tbody>
</table>

Table 4.4: Correlation values between each assignment and the Moodle access corresponding to the time assignments and Moodle access during those assignments.

Table 4.4 shows the correlation values between assignment grades and their corresponding access during the time of the assignment. From these results, it appeared that while assignment 1 (A1) and assignment 2 (A2) had almost no correlation to LMS access \( r_s(75) = 0.04, p > 0.05 \) and \( r_s(75) = 0.05, p > 0.05 \) respectively), assignment 3 (A3) had a slight positive correlation \( r_s(75) = 0.25, p < 0.034 \).

The amount of access during the earlier assignments did not seem to have an impact on how well the participant did on that assignment, whereas during assignment three more access indicated better grades. There are many possible causes for this change. Perhaps by the time A3 came around students had learned which sections of the website were the most useful, making Moodle access more beneficial to their grades. Or maybe students did not utilize the forums during the earlier assignments, but had become comfortable enough to ask or answer questions during A3. Another possibility is that the Progress Tracker did make a difference (PT views were positively correlated with A3 grades \( r_s(75) = 0.24, p < 0.037 \) and students who saw their progress more often were more motivated to do well on the assignment.

The unexpected correlations found here raise questions about how students interact with the e-learning system. A discussion about what these results
suggest and how they could be used to improve the Progress Tracker plug-in will follow in Section 4.6.

4.4 Changes in Variables Indicating Engagement

In Section 3.4 Question 1 asked: “Can students’ engagement in an e-learning environment be improved by providing comparative visualizations of the students’ progress?” In order to determine if the plug-in positively influenced engagement in the class, several variables were compared across the three assignments. Variables were selected based on their success at indicating engagement in previous studies (see Section 2.4 of the literature review). The following types of variables were examined because they were similar to those used in previous studies to measure engagement (see Section 2.4):

- Checklist utilizations (both number of updates and completion percentage),
- Assignment start times (when students first viewed the assignment specification on Moodle)
- Commits with Git (the number of recorded changes while working on the assignment)
- Assignment submission times (the time at which students submitted their final version of their assignment via Git and their collaboration statement via Moodle)
- General access to Moodle

In e-learning, engagement is strongly linked to success and vice versa (Carini et al., 2006). Therefore, if these variables could indicate success (achievement of the students’ goals) in this particular data set then they were likely to indicate engagement as well. To verify that the chosen variables were correlated
with the participants’ success, they were first compared between the group of participants who met the goal they had set for themselves (successful) against those who did not meet their goal (unsuccessful). Table 4.5 shows the variable averages and medians for each group. (Git submits were measured as how many hours before or after the deadline the submission was made. Therefore positive numbers indicate hours early and negative numbers indicate hours late.) Every one of the variables showed the more favourable measurements (more updates and commits, higher completion and access, and earlier submits and start times) in the successful group. This shows that these variables were able to indicate success in this data set, and based on previous studies, are likely to indicate engagement as well.

Table 4.5: A3 variable averages and medians for participants who were successful (met their goal) and those who were not.

<table>
<thead>
<tr>
<th>A3 Variable</th>
<th>Successful</th>
<th></th>
<th></th>
<th>Unsuccessful</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Median</td>
<td>Average</td>
<td>Median</td>
<td>Average</td>
<td>Median</td>
</tr>
<tr>
<td>Checklist updates</td>
<td>15.33</td>
<td>15</td>
<td>15.14</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checklist completion</td>
<td>75.33</td>
<td>95</td>
<td>59.11</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commits</td>
<td>10.47</td>
<td>10</td>
<td>7.60</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Git submit</td>
<td>7.10</td>
<td>9</td>
<td>-4.11</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start time</td>
<td>149.98</td>
<td>30.25</td>
<td>202.39</td>
<td>125.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moodle submit</td>
<td>35.69</td>
<td>10</td>
<td>9.95</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access</td>
<td>26.39</td>
<td>16.4</td>
<td>14.91</td>
<td>11.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data from A1 and A2 were treated as control measurements for the engagement level before the addition of the plug-in. Data from A1 and A2 were compared to the data recorded during A3, when the Progress Tracker was active. Table 4.6 contains the averages and medians for engagement-indicative variables across the three assignments.
Table 4.6: Variable averages and medians across assignments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Assignment 1</th>
<th>Assignment 2</th>
<th>Assignment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Median</td>
<td>Average</td>
</tr>
<tr>
<td>Checklist completion</td>
<td>75.83</td>
<td>95</td>
<td>64.28</td>
</tr>
<tr>
<td>Commits</td>
<td>4.64</td>
<td>2</td>
<td>6.33</td>
</tr>
<tr>
<td>Git submit</td>
<td>14.62</td>
<td>13</td>
<td>-29.29</td>
</tr>
<tr>
<td>Start time</td>
<td>111.48</td>
<td>44</td>
<td>258.83</td>
</tr>
<tr>
<td>Moodle submit</td>
<td>0.58</td>
<td>10</td>
<td>13.83</td>
</tr>
<tr>
<td>Access per day</td>
<td>14.99</td>
<td>11</td>
<td>9.72</td>
</tr>
</tbody>
</table>

The Friedman test was used to check for significant differences between the measurement of each variable during A1, A2, and A3. This test detects differences in treatments across multiple test attempts where measurements are non-normal and non-independent (Mendenhall & Beaver, 1994), therefore it was chosen as the most suitable for this application. In addition, the Bonferroni correction was used to account for the “multiple testing” problem. Basically, when multiple pair-wise comparisons are done the chance of type I (false positives) increases (Bland & Altman, 1995). The Bonferroni adjustment modifies the level of significance from $p < \alpha$ to $p < \alpha/n$ (where $n$ is the number of comparisons) in order to correct for the higher chance of type I error (Bland & Altman, 1995). In this case, the Bonferroni adjustment resulted in significance levels between $p < 0.00067$ to $p < 0.00083$ for these seven variables. The results of the Friedman test are displayed in table 4.7. Results highlighted in green indicate those which deemed the variable significantly different between the two assignments being compared.

The Friedman test in conjunction with the Bonferroni adjustment indicated that the checklist completion and Git submits on A3 was significantly different from that of A1. It also showed that A3 commits were significantly different...
from both A1 and A2. There were no significant differences across the other variables according to the Friedman test.

Graphs of the averages of these variables along with their standard error are displayed in Figure 4.7. Note that the standard error is actually displayed as $2 \times stderr$ to represent the approximate 95% confidence interval. The average checklist completion percentage on A3 was lower than that of A1 by 14.27% (Figure 4.7a). The average Git submission time for A3 was 22.53 hours later than those of A1, with the A3 average submission being 7.91 hours late. However, the Git submission measure had a very high standard error (Figure 4.7b). The number of commits increased on A3 from the counts on both A1 and A2. A3 had an average of 7.96 commits whereas A1 and A2 had 4.64 and 6.33 respectively (Figure 4.7c). Last of all, there was a significant difference in the average Moodle accesses per day from A2 to A3 (increasing from 9.72 to 14.06), however the change from A1 to A3 was not significant (Figure 4.7d).

The median measures of the variables are shown in Figure 4.8. As seen in Figure 4.8a, the median checklist completion percentage dropped much more noticeably than the average from A1 to A3 (a decrease of 25%). The median values for Git submission times (Figure 4.8b) were not as drastically different

<table>
<thead>
<tr>
<th>Variable</th>
<th>Friedman test results</th>
<th>Friedman test results</th>
<th>Friedman test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1 vs. A2</td>
<td>A1 vs. A3</td>
<td>A2 vs. A3</td>
</tr>
<tr>
<td>Checklist updates</td>
<td>$X^2(1) = 0.2424, p = 0.6225$</td>
<td>$X^2(1) = 0.9143, p = 0.339$</td>
<td>$X^2(1) = 1.1739, p = 0.2786$</td>
</tr>
<tr>
<td>Checklist completion</td>
<td>$X^2(1) = 5.1212, p = 0.02364$</td>
<td>$X^2(1) = 12.3023, p = 0.00045$</td>
<td>$X^2(1) = 0.6098, p = 0.4349$</td>
</tr>
<tr>
<td>Commits</td>
<td>$X^2(1) = 3.6885, p = 0.05479$</td>
<td>$X^2(1) = 36.6338, p = 1.425e^{-9}$</td>
<td>$X^2(1) = 27.597, p = 1.494e^{-7}$</td>
</tr>
<tr>
<td>Git submit</td>
<td>$X^2(1) = 12.6452, p = 0.00038$</td>
<td>$X^2(1) = 28.4462, p = 9.634e^{-8}$</td>
<td>$X^2(1) = 0, p = 1.00000$</td>
</tr>
<tr>
<td>Start time</td>
<td>$X^2(1) = 4.3784, p = 0.0364$</td>
<td>$X^2(1) = 9.72, p = 0.00182$</td>
<td>$X^2(1) = 0.3333, p = 0.5637$</td>
</tr>
<tr>
<td>Moodle submit</td>
<td>$X^2(1) = 2.0571, p = 0.1515$</td>
<td>$X^2(1) = 1.3729, p = 0.2413$</td>
<td>$X^2(1) = 1.6667, p = 0.1967$</td>
</tr>
<tr>
<td>Access per day</td>
<td>$X^2(1) = 49.6133, p = 1.87e^{-12}$</td>
<td>$X^2(1) = 8.3333, p = 0.00389$</td>
<td>$X^2(1) = 23.8378, p = 1.048e^{-6}$</td>
</tr>
</tbody>
</table>

Table 4.7: Results of the Friedman test on variables compared between assignments.
Figure 4.7: Averages of variables deemed significantly different across assignments.

as the averages, only 13 hours early on A1 and 4 hours early on A3. However, the increase in median Git commits from A1 and A2 to that of A3 was higher
than the average. Median Git commits went from 2 and 3 on A1 and A2 to 7 commits on A3 (Figure 4.8c). With the averages there was a significant difference in Moodle accesses from A2 to A3, however when looking at the medians (Figure 4.8d) the increase is much less noticeable.

A similar comparison was done between the participants who set a goal through the Progress Tracker and those who did not. The same engagement-indicating attributes for assignment three were compared between these two groups. The differences in these were clearer than those between assignments one, two and three. For every one of the variables being looked at, the group who had set goals had the more favourable measurement. Table 4.8 shows the average and median across the two groups for each of the variables reviewed.

<table>
<thead>
<tr>
<th>A3 Variable</th>
<th>Goal set</th>
<th></th>
<th>Goal not set</th>
<th></th>
<th>Wilcoxon test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Median</td>
<td>Average</td>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>Checklist updates</td>
<td>15.19</td>
<td>14</td>
<td>11.41</td>
<td>7</td>
<td>W = 649.5, p = 0.04798</td>
</tr>
<tr>
<td>Checklist completion</td>
<td>65.68</td>
<td>83</td>
<td>58.57</td>
<td>65</td>
<td>W = 498.5, p = 0.30200</td>
</tr>
<tr>
<td>Commits</td>
<td>8.34</td>
<td>8</td>
<td>6.65</td>
<td>6</td>
<td>W = 594.5, p = 0.20010</td>
</tr>
<tr>
<td>Git submit</td>
<td>4.86</td>
<td>5</td>
<td>–52.61</td>
<td>1</td>
<td>W = 576.0, p = 0.08384</td>
</tr>
<tr>
<td>Start time</td>
<td>185.42</td>
<td>80</td>
<td>306.84</td>
<td>294</td>
<td>W = 318.0, p = 0.02708</td>
</tr>
<tr>
<td>Moodle submit</td>
<td>16.17</td>
<td>8</td>
<td>9.27</td>
<td>7</td>
<td>W = 454.0, p = 0.25080</td>
</tr>
<tr>
<td>Access</td>
<td>16.30</td>
<td>11</td>
<td>6.41</td>
<td>5</td>
<td>W = 812.0, p = 0.00006</td>
</tr>
</tbody>
</table>

Table 4.8: A3 variable averages and medians for goal set vs. no goal set

With this data, the Wilcoxon rank-sum test (also known as the Mann-Whitney U test) was used to look for significant differences between the two groups. The Wilcoxon test is a non-parametric test of a null hypothesis between two sample populations where the observations of both populations are independent of each other (Mendenhall & Beaver, 1994). It was chosen over the T-Test because it is more effective on non-normal distributions and allows for ordinal data where the spacing between adjacent values cannot be assumed to be constant.
Figure 4.8: Medians of variables deemed significantly different across assignments.

When applying the Wilcoxon test to this data, it was assumed that each participant’s measurements were independent of the other participants. The
test was run as a two-tailed test with the null hypothesis that the median difference between the goal and no goal groups was zero. The typical significance level of $p < 0.05$ was used to determine which variables were different between the two groups. The last two columns of table 4.8 give the test statistic ($U$) and the $P$-value for each attribute.

Although all of the variables had favourable measurements for the group who had set goals, the Wilcoxon test showed only three as being significantly different between the two groups (see Figure 4.8; the number of checklist updates, assignment start time, and access during the assignment. The other four variables were not different enough to be of significance. Graphs of the averages of the variables deemed significantly different are displayed in Figure 4.9. Note that the standard error for averages is displayed as $2 \times stderr$ to represent the approximate 95% confidence interval. The median results are shown in Figure 4.10.

In Figure 4.9a it can be seen that the group who set goals had more checklist updates (15.19 for those with goals and 11.41 for those without). The assignment three start times were much better for the goal group, with an average start time 121.42 hours earlier than the no goal group (Figure 4.9b). Finally, Moodle access during assignment three had the highest measured difference between the goal and no goal group, with $p = 0.00006$. Those who had set goals averaged 16.3 accesses per day, and those without a goal averaged only 6.41 accesses per day (Figure 4.9c).

The median number of checklist updates during assignment three for the goal setting group (14) was double that of the group without goals set (7) (Figure 4.10a), a more drastic difference than seen with the average measurements. The median measurements of the assignment three start times (shown in Figure 4.10b) also had a higher degree of difference between the groups than the
average measurements. The goal group had a median start time of 80 hours after the time the assignment was posted, whereas the no goal group’s median start time wasn’t until 294 hours after the post time (a difference of 214 hours). Moodle access showed a larger difference with the average measurements, but the median access per day shows the same trend (Figure 4.10c). The goal group had a 11 median daily accesses and the no goal group had 5 median daily accesses.

4.5 Clustering

To further explore the connections between the participants’ results in the class and their interactions with the LMS, the data was further examined through clustering. Weka’s K-means clustering algorithm was applied to the numerical data set with the parameters discussed in Section 3.2.4. The numerical data set
was chosen over the letter grade data set because K-means is intended for continuous numeric values rather than categorical ones. Weka’s implementation of K-means is capable of handling categorical and binary data types, however its error output would not make sense because the mean cannot be calculated for categorical attributes. In addition, because the letter grades are not ordinal from Weka’s point of view (i.e. Weka does not know that an A is close to an A+) the clustering of grade data would not be as accurate as it would with numerical values (i.e. Weka does know that an 87% is close to a 95%).

Some unnecessary attributes were removed for the clustering of the data. Participant IDs were excluded since they were only an identifier. The accumulated grade attribute and calculated averages for assignments, lectures, and

Figure 4.10: Medians of variables deemed significantly different between goal/no goal groups.
quizzes were removed since they are entirely comprised of other graded attributes. Lastly, individual lecture attendance data was removed in favour of the single “total attended” attribute to avoid including so many binary values.

It was also necessary to determine the best number of clusters to use to analyze the data. Increasing the number of clusters without penalty will reduce the error in the resulting clusters, to the extreme case of zero, where every data point is in its own cluster. However, this would provide no information since the data is no longer being grouped according to similarities. Ideally, the number of clusters should be high enough as to reduce the measured error, but not so high that the size of the clusters become too small. Therefore, the algorithm was run four times, with each run increasing the number of clusters from two to five, and the resulting cluster sizes and errors were compared. Table 4.9 shows the mean squared error (the average error across each of the clusters) as well as the size of each cluster created.

<table>
<thead>
<tr>
<th>Number of clusters</th>
<th>Mean square error</th>
<th>Cluster size</th>
<th>Cluster1</th>
<th>Cluster2</th>
<th>Cluster3</th>
<th>Cluster4</th>
<th>Cluster5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>338.55</td>
<td></td>
<td>32 (43%)</td>
<td>43 (57%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>330.13</td>
<td></td>
<td>19 (25%)</td>
<td>26 (35%)</td>
<td>30 (40%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>333.90</td>
<td></td>
<td>9 (12%)</td>
<td>16 (21%)</td>
<td>43 (57%)</td>
<td>7 (9%)</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>302.99</td>
<td></td>
<td>8 (11%)</td>
<td>49 (65%)</td>
<td>7 (9%)</td>
<td>6 (8%)</td>
<td>5 (7%)</td>
</tr>
</tbody>
</table>

Table 4.9: Results from varying the number of clusters with K-means algorithm.

From table 4.9 it can be seen that using three clusters gave the best results for this data set. When clustered into three groups, the cluster sizes remained evenly distributed (19, 26, and 30 participants across the three clusters). The error has also been reduced from that of the two-cluster results. When the number of clusters was increased to four and five, the cluster sizes became much more uneven, with a single cluster containing the majority of participants and the others having only small populations. The error also increased from
the three-cluster results to the four-cluster results, making it clear that three clusters were favourable in this case.

Figure 4.11: Trends in course graded components for each cluster.

Figure 4.11 displays the trends among clusters for any graded attribute. From looking at the “final” attribute cluster two (C2) had the highest final grades, averaging at 87%, cluster one (C1) had an average final grade of 77%, and cluster three (C3) had the lowest grades averaging at 63%. Throughout the graded attributes for the course (all those that make up the final grade from “A1” to “Exam”) C2 remains at the top and C3 at the bottom, with C1 jumping between the grades of the other two clusters. On A2, LE1, and Clicker, C1 is on par with the higher grades of C2. However on A3 and Q1 C1 drops to meet the lower grades of C3. It is also interesting to note how large of a difference there is between the A3 grade for C2 and that of the other two clusters.

Figure 4.12 displays the average results of the SCEQ for each cluster. As with the correlation results discussed in Section 4.3, the overall SCEQ grades
are not reflective of the grades received on the graded components in the course. All three clusters have a similar total SCEQ score, but vary across the four subcategories within. C1 had the highest “skills” scores even though they did not display the highest grades in the course. C2 and C3 received almost the
same “skills” scores despite the fact that their course grades were so contrasting. Note that the SCEQ “skills” score is a self-reported measure of a student’s study habits rather than their skills in the course content. The difference in the “emotional” and “participation” areas between clusters was small, however it is interesting that C3, the group with the lowest grades had the highest “emotional” scores, and that C2, the group with the highest grades, had the lowest “participation” scores. The “performance” measure appeared to be the only one of the four SCEQ areas that was clearly reflective of grades earned in the course.

Figure 4.13 shows the checklist completion level for each cluster on each assignment. Throughout all three assignments C3 completed the least of their checklists, which correlated with their grades on those assignments. During assignments one and two, C1 completed more of their checklists than C2, however on assignment three, C2 completed more and C1 completed less. The resulting assignment three grade for C1 was also considerably lower than that of the other two assignments, perhaps due in part to their drop in the fulfillment of their checklist during that assignment.

Moodle access for each cluster throughout the semester is displayed in Figure 4.14. The time spans along the x-axis should be fairly self-evident. “A1” indicates the average access per day from the time assignment one was posted to the time it was due, “A1-A2” indicates access between the due date of assignment one and the post date of assignment two, and so on. “Exam to grade post” indicates access from after the final exam to the time at which the university posted official grades.

As expected, C2, which had the highest grades out of the three clusters, also had the most Moodle access overall. Interestingly, C2’s access dropped below that of C1 during assignment two, which was also the assignment that C1 did
Figure 4.14: Trends in Moodle Access for each cluster throughout the semester. The best on, matching the grade of C2. This indicates that the amount of Moodle access does play a role in the success of students on their assignments.

From this graph it can also be seen that C1 tends to have higher access during assignments, and lower access between assignments. The other two clusters, C2 and C3, stay more steady in their access through the semester.

Figure 4.15: Attendance, forum, and Progress Tracker view data across clusters.
Attendance data is illustrated in Figure 4.15a. Surprisingly lecture attendance results did not follow the same trend as grades and Moodle access. C1, the mid-grade cluster, had the highest lecture attendance, followed by C2, then C3. Forum views and Progress Tracker views (shown in Figures 4.15b and 4.15c respectively) did follow the same pattern as grades and Moodle access. C1 had moderate views, C2 had the highest number of views, and C2 viewed forums and the Progress Tracker the least.

![Comparison of variables across clusters](image)

(a) Average number of hours between assignment post and when the participant viewed the assignment  
(b) Average number of checklist updates for assignments  
(c) Average Git commits  
(d) Average Git submission  
(e) Average Moodle submission

Figure 4.16: Comparison of variables across clusters

Figure 4.16a shows how soon participants from each cluster started their assignment after the assignment description had been posted on Moodle. Here a lower number is favourable, it means the participants started sooner rather
than later. C1 was first to start their assignments, but with each subsequent assignment they started later than with the assignment before. As expected, C3 started the latest, and C2 was in between, starting A1 around the same time as C3 and A3 at the same time as C1.

The number of checklist updates during each assignment followed the same trend as assignment start times. From Figure 4.16b it can be seen that C1 had the most checklist updates, followed by C2, then C3.

The number of Git commits (Figure 4.16c) did not have a clear leader like the other measures. However, each cluster showed different trends between the Git commits on each assignment. C1 showed a steady increase in commits between assignments one and three, C2 showed a steep increase from commits on assignment two to assignment three, and C3 showed an opposite decrease in commits from assignment two to assignment three. This steep increase for C2 reflects the much higher grade achieved on assignment three compared to the other two clusters. Meanwhile the decrease for C3 seems linked to their much lower grade on assignment three compared to the other two assignment grades for that cluster.

Assignment submissions, both the Git code submission and the Moodle collaboration statement (illustrated in Figures 4.16d and 4.16e respectively), show C3 as submitting later than the other two clusters. C2 had the earliest Git submits, but not the earliest Moodle submissions. This suggests perhaps that they were more confident in their code, not requiring later adjustments during the remark period.

By looking at each piece of this information as part of a larger puzzle, hypothetical profiles can be created in an attempt to generalize and explain the similarities among participants in each cluster.
Cluster 1 had mediocre grade averages, yet had the highest SCEQ skills score, lecture attendance, and clicker grades. They had the earliest assignment start times, the most checklist updates, and the highest checklist completion on assignments one and two. C1 also had steadily increasing numbers of Git commits and upward spikes in their access to Moodle during assignment times. From this information, one could speculate that these students are hard workers with good study habits and who like to seek opportunities to help themselves learn. They may struggle to comprehend and apply the material, but they attempt to remedy this by starting earlier and keeping themselves organized.

Cluster 2 achieved the highest grades, leading by a large margin on assignment three. They had high SCEQ performance measures, but low skills and participation measures, as well as lower lecture attendance than C1. Their clicker grade was the only graded component in which they were not the highest. C2 did however have the highest forum views, Progress Tracker views, and overall access throughout the semester. They also had much higher Moodle access during the times between assignments than the other two clusters. From this data it appears that participants in C2 were quite confident and capable in their coding abilities since they received consistently high grades throughout the semester. They had the best grades by far on assignment three, which was the most challenging of the three, suggesting that their coding abilities were superior to the other clusters. The fact that their lecture attendance, study skills, and in-class participation were lower suggests that they were confident enough that they did not feel as though they needed to access help to keep their grades high. The high and steady access could mean that they found answers to their own questions through the course website content and the forums. The high forum views could also mean they were replying to other students’ questions as well as looking for answers to their own questions and
that they may learn via peers rather than traditional classroom methods. It seems these participants had a natural aptitude for the content in this course and may benefit from a more fast paced and challenging learning environment.

Cluster 3 had the lowest grades of the three. They also had the lowest lecture attendance, performance SCEQ score, checklist updates and completion, forum views, Progress Tracker views, and Moodle access during assignments. C3 also started and submitted assignments the latest. The average Git submission time for this cluster was later than the due date for all three assignments, suggesting that many of these participants either submitted late in the first place or required re-grades on their assignments. C3 did however have the highest emotional and participation scores on the SCEQ. This group appeared to be what are commonly referred to as the “slackers”. They did not attend many lectures and attained low grades, with seemingly little effort to raise their marks. Perhaps starting earlier on assignments or attempting to answer their own questions via the content on Moodle or the forums would have helped but they were not motivated to do so. C3’s high scores on the emotional and participation sections of the SCEQ were surprising. Since the participation scores did not reflect their actual lecture attendance or clicker scores, they may have simply scored themselves higher on the questionnaire to compensate for their lack of knowledge and hard work. Their high emotional score suggests that perhaps they were interested in the content but had difficulty with the learning format used for the course. These students could have also been embarrassed about their lack of understanding, and therefore did poorly because they did not want to ask for help or clarification.
4.6 Discussion

In Section 3.4 two questions were posed for examination. Each of these questions was answered through one or more methods of analysis performed on the data.

Question 1 asked: “Can students’ engagement in an e-learning environment be improved by providing comparative visualizations of the students’ progress?” In order to answer this query, variables that have been shown to indicate engagement in similar studies were compared across the three assignments in the course, then again across the group of participants who set goals through the Progress Tracker and those who did not. The results of this analysis are shown in Section 4.4. Favourable measures of these engagement indicative variables during assignment three or in the group who used the Progress Tracker would suggest that the introduction of the Progress Tracker for assignment three may have increased engagement.

When comparing variables across assignments, changes in the variables from A1 and A2 to that of A3 were investigated. There were significant changes in the completion percentage of checklists, the number of Git commits, Git submission times, and the amount of Moodle access from A1 and/or A2 to A3 (results are displayed in table 4.6). Moodle access during A3 was higher than that during A2, but was still lower than A1. Checklist completion and Git submits were both lowest during A3. The number of Git commits during the assignment was the only variable to see overall improvement during A3. However, A3 introduced the element of “checkpoints” where the students could submit their assignment early to discover its flaws before the actual due date. This is likely the reason for the increase in Git commits during the final assignment. Therefore, from this comparison of data it does not appear that the Progress Tracker contributed to increased engagement.
The comparison of the same variables between the group of participants who used the Progress Tracker to set goals and those who did not, yielded more promising results. Here there were significant changes in the number of checklist updates during A3, the time participants started A3, and the amount of Moodle access during A3 (results are displayed in table 4.8). In this comparison, the group that had set goals had more favourable measurements for all variables (including those that did not show significant differences). They updated their checklists more often, started their assignment earlier, and accessed Moodle much more often than the group without goals. This either indicates that the Progress Tracker had a positive influence on engagement for those who used it, or that only the more engaged participants chose to use the PT. Additional research would need to be performed in order to determine which scenario occurred.

Question 2 asked: “Do students’ believe that the visualization helped improve their motivation in the e-learning environment?” Information about the participants’ feelings towards the Progress Tracker were collected through the PTFQ. The PTFQ results are presented in Section 4.2 and the responses to each question asked can be seen in Figure 4.4. For every one of the 10 questions the majority of the responses were positive.

The two statements with the most positive responses confirmed that participants enjoyed the progress monitoring for their goals and being able to compare their progress with that of their peers. These were both a main objective in creating the Progress Tracker. Only one participant disliked the look of the plug-in and most participants agreed that they would use the plug-in in other courses if they had the option. Judging only from the PTFQ responses, it seems that the self-progress towards the participants’ goals was more motivating than knowing how the rest of the class was doing. The statement:
“The information provided through the plug-in made me more competitive”, had the most negative responses, suggesting that this aspect of the plug-in was the least successful according to the participants.

The feedback from the PTFQ was able to provide levels of success for the Progress Tracker objectives listed in Section 3.3.1:

1. To provide students with a tool for goal-setting that also keeps them aware of their progress towards that goal
2. To remind students of their progress on course content and how much time remained before the due date of that content
3. To allow students to compare their progress to that of their peers, and possibly introduce a competitive aspect into the e-learning system
4. To create an aesthetically pleasing and intuitive interface that students enjoy interacting with

Judging from the participants’ feedback on the PTFQ, objective one was successful. 84% of respondents stated that they “liked being able to see my progress towards my goal”. 79% said that “being reminded of my own progress helped motivate me”, which provides partial feedback for objective 2. However, there were no statements on the PTFQ about the effectiveness of due date reminders. The statement “I liked being able to compare my overall grade to the rest of the class” had the most positive responses (89%) which shows that the first part of objective 3 was appreciated. However, only 37% felt that the plug-in made them more competitive. Objective 4 was partially successful. Only one participant stated that they disliked the look of the plug-in and although there were no statements specifically regarding the intuitiveness of the interface, two participants stated that the Progress Tracker was not clear or obvious enough in their short answer responses.
The PTFQ feedback provided encouraging information about the satisfaction and effectiveness of the plug-in. Although not all responses were positive, they helped to identify the problem areas. Some of the participants’ short answer responses also gave constructive criticism for how the plug-in could be improved. These will be discussed in Section 5.1.

Question three asked: “What knowledge can be acquired from the students’ data in order to better induce engagement through information visualization?” Analysis of the data through correlation and clustering produced a lot of information about participant behaviours and how they interacted with the LMS. This information provides some insights for how the plug-in could be improved for use in future studies.

Results from cluster analysis of the data yielded three types of students:

1. Students who put a lot of effort into studying, yet receive mostly mediocre grades despite their best attempts
2. Students who do quite well and receive high grades, but do so through a natural aptitude and seemingly without challenge
3. Students who seem to lack the motivation or skills to do well, and therefore end up with low grades

From observing the characteristics of each cluster, it appeared that the correlations found in the data for all participants were also very applicable to the clusters. Interestingly enough, the SCEQ areas seemed to be very suggestive in which cluster a participant belonged. High emotional and performance scores were seen primarily in the high-grade cluster, high skills scores were seen in the middle-grade cluster, and high participation scores were seen in the low-grade cluster.
Judging from the hypothetical profiles of the participants outlined in Section 4.5 and the trends seen in the correlation matrix discussed in Section 4.3, these three types of students would benefit from the plug-in in different ways. The high-grade cluster did not attend lectures as often as others, however they did have the most forum access, so perhaps they learn via peers rather than the traditional methods. They may seek more fast-paced and challenging learning environments with students like themselves. If they often confer with skilled peers, then perhaps they are also interested in how well the stronger peers are doing on their assignments. This group is probably proud of their skills and may be more likely than others to be competitive.

The middle-grade cluster accessed the LMS more during assignments and they used the checklists the most often. It’s fairly clear that they require more effort to do well, and may appreciate the step-by-step form of instruction (like the checklists). This group also started and submitted their assignments early, so perhaps they liked having reminders of important dates. From the level of effort they appear to put into their studies, it seems that this group would benefit from the goal-setting and monitoring aspect of the plug-in.

The low-grade cluster put in seemingly little effort toward attaining good marks. This could be because they were not motivated to do so, they could have had difficulty learning in the format used for CIS*2500, or perhaps they were too shy to ask for help when they needed it. These students started and submitted assignments later than the other groups. Perhaps the most beneficial feature of the plug-in for this group would be the class assignment progress monitoring and important date reminders. Being aware of when their classmates start the assignment might encourage them to start earlier and therefore be able to finish their assignments in order to hand them in by the deadline.
As is, the plug-in seems to have been more useful for the high-grade and middle-grade clusters. The low grade cluster had an average of only 3.6 Progress Tracker views, indicating that it did not contain information that they felt was useful to them. There are many possibilities for how to improve the plug-in for the types of students in each cluster. A discussion of these ideas will follow in Section 5.1.
Chapter 5

Summary and Future Work

Progress Tracker, a Moodle plug-in was created with the aim of increasing student engagement through data visualization of self-progress information, goal setting, and peer comparison. The plug-in was deployed in an intermediate programming course for the final third of the semester. Data from 75 students was used to analyze the differences in engagement caused by the introduction of the plug-in.

Analysis of engagement-indicative variables between the participants who used the plug-in and set a goal against those who did not set a goal showed significant improvements in three of the seven variables examined. All seven areas had favourable measurements in the goal-setting group, however the other four were not statistically significant. Students who set a goal using the Progress Tracker had 33% more checklist updates, 40% earlier start times, and 154% higher Moodle access. These results show that those who set goals for themselves in the Progress Tracker were more active, possibly indicating higher engagement. It cannot be concluded that these improvements were caused by the PT, although that is a possibility. It could be simply that the students who
were naturally more engaged used the Progress Tracker and the less engaged students did not.

Examination of the same variables was done between A1 and A2 (before the PT) against those measured during A3 (with the PT). Here 4 of the 7 variables showed significant differences during A3, some favourable and some unfavourable. The number of recorded changes by the students (via Git commits) increased 72% from A1 to A3 and 26% from A2 to A3. Git submits for A3 were 21 hours earlier on average than A2 but 22 hours later than A1. Moodle accesses increased 45% from A2 to A3 but decreased 6% from A1. Checklist completion decreased by 4% from A2 to A3 and 19% from A1 to A3. The number of Git commits improved compared to both A1 and A2, however submission times and Moodle access improved only compared to A2. Given that these were the only variables out of the 6 examined (see Section 4.4) that showed significant improvement, it cannot be concluded that the Progress Tracker was related to an increase in engagement in the class.

Participants were also asked for feedback on the Progress Tracker through the use of a questionnaire. The responses collected were very positive, indicating that most participants who used the Progress Tracker and completed the questionnaire, enjoyed the plug-in and believed that it motivated them. 5 of the responses also provided suggestions for ways in which the plug-in could be improved.

Participants’ feedback indicates that they felt the plug-in was successful with respect to the objectives that were laid out in Section 3.3.1. Although the analysis of the data did not show concrete evidence of this the collective results are encouraging, and with some improvements the plug-in could be made more effective.
5.1 Future Work

The work done in this study could be expanded on in many ways. The possibilities for improvement and future development of the study will be presented twofold. First, suggestions for improvement and additions to the plug-in will be discussed, then changes for future evaluation of the plug-in through experimentation will be recommended.

For the purpose of the research, the plug-in was kept quite simple in its content and operation. The ability to customize views and content, so that students can centralize the information they feel is most important and see that information in a way that is preferred by them, could be a worthwhile addition in future editions. There are many opportunities for additional features and new content that students might find interesting, and adding the ability for students to choose which information is displayed to reduce clutter and allow for personalization.

Ideas for some potentially useful additions to both Moodle and the plug-in were spurred by the results gathered through clustering of the data. A point or voting system could be introduced to the forums. Students could up-vote or down-vote responses, both making it easier for other students to identify the useful posts and for providing incentives for students to give helpful responses. The high-grade cluster might be more inclined to share their knowledge on the forums if they were rewarded with a higher status for doing so. It may also introduce a more competitive atmosphere. The middle-grade and low-grade clusters would benefit because they would receive more helpful answers to their questions and be able to find information much more easily. In addition, this would provide useful data for evaluating students’ engagement in future studies.
The low-grade cluster displayed much lower forum views than the other two clusters. These students may not have found the forums helpful, which may be remedied by the voting-system just discussed, or they did not want to post their questions. One probable cause for students not asking for help when they need it might simply be that they are too embarrassed in their lack of knowledge to do so. Perhaps the addition of an anonymous forum could help these students. They may be more inclined to ask questions they feel are “stupid” if their peers don’t know who asked the question.

The feedback gathered with the PTFQ indicated that the meaning of the graphs could be made clearer. Some expressed concern that it was not entirely obvious how the plug-in displays progress and what the progress means. Other ways to display the progress data should be explored in order to make the visualizations more obvious. Usability testing should also be done to ensure that the information is easily understood before experimentation.

It may also be worthwhile to implement a checklist-type feature into the plug-in directly. In this study students had to use the existing checklist plug-in for their progress data on assignments to be displayed through the PT. If the checklist for completing each portion of the assignment was located in the PT, then students would be reminded of their progress each time they went to view their checklist. They would also be more likely to update their checklist when viewing the Progress Tracker since it would be more convenient to do so. The plug-in should also be made so that students can update their goals from the more convenient block portion rather than requiring them to navigate to the module. Perhaps more students would set and maintain goals if they could set them from any course main page.

For this study the Progress Tracker was deployed in an actual class to evaluate its effectiveness. Because the experiment was not carried out in a
controlled environment, several limitations were present. The measurement of engagement-indicative variables during assignment three (when the Progress Tracker was available) were compared to assignments one and two. Ideally, these three assignments would be identical in terms of structure so that assignments one and two would act as reliable control measures. However, these assignments were part of an actual course and designed by the instructor, therefore the researcher had no control over their structure. In the case of CIS*2500, a series of checkpoints (where the students could submit their code ahead of the due date to verify that it worked correctly) were present for assignment three only. These checkpoints would have altered how early students started their assignments, how many times they committed their work to Git, and their Moodle access patterns. If the assignments, or tasks, could be made identical in structure then one could be used for a reliable control measure while the other tests the effectiveness of the plug-in. This way, any changes in the engagement-indicative variables between the two could be much more accurately attributed to the plug-in.

Carrying out the study in a controlled environment would also simplify other areas of the experiment. According to the REB, all students must be provided with the same opportunities, therefore the plug-in had to be available to all of them. If the experiment were done outside of an actual course with participants rather than students, one half could be given the plug-in while the other half would act as control subjects. This would provide reliable measurements for evaluating the plug-in. Another upside would be that all of the participants who completed the study would complete the feedback questionnaire. In this study, due to the timing of recruitment, only 22 participants completed the PTFQ out of the 75 total participants, and of those 22, only 19
actually used the plug-in. It would have been beneficial to know the opinions of the other 53 participants.

In future experimentation, the collection additional supportive variables in the data set may also be of use. In this study the length of logged Moodle access was not analyzed, only the time at which the start of the access was recorded. It may be useful to know how long a participant accesses each Moodle module. It would also be helpful to know whether the participant was active during the recorded access or whether the page was simply left open. Although Moodle does not have that ability at this time, there are studies being done on recognizing a user’s activity level through clicks and scrolling (Atterer, Wnuk, & Schmidt, 2006).

This study did not look at participants’ forum posts, only the logs of when they viewed the forums. For future studies, it would be useful to identify the participants that actively post questions and answers on the forums versus those that were “lurkers” and only read existing posts. This data is already available through Moodle and could be added to the data set. It would also be interesting to identify those who post questions and those who post answers, those who answer a post versus making the initial post, the size of posts, and the length of threads. Semantic analysis may be able to discern the difference by looking at the content of the participants’ posts.

The clustering analysis in this study was done using Weka, and was therefore subject to Weka’s methods and limitations. It would have been interesting to know what ratio of participants in each cluster had met their goals. But since this was a binary variable, it could not be included in the data used for clustering, and even if it could have been, it would have influenced the results. In the future, the analysis should be done using a method whereby participant IDs can be kept intact through the clustering process. That way
trends can be identified afterward by cross-referencing the participant IDs in each cluster with variables from the original data set that were not included in the clustering.

5.2 Conclusion

This study has shown encouraging results by using a Moodle plug-in to boost student engagement through a combination of self-progress information, goal setting, and peer comparison. The results were not concrete enough to say without a doubt that engagement was improved, however they showed promise in several areas. Feedback from the participants was very positive, with many stating that they felt their motivation was increased by the plug-in. With a few small changes to the plug-in and a controlled method to evaluate its effectiveness, more obvious increases in engagement would likely be seen. If the plug-in could contribute to even a small increase in student engagement and motivation, it would be an important step forward in the design and effectiveness of Learning Management Systems.
Appendix A

Supplementary Information

Table A.1: Abbreviations used throughout thesis

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1, A2, A3</td>
<td>Assignment 1, 2, and 3 respectively</td>
</tr>
<tr>
<td>C1, C2, C3</td>
<td>Clusters 1, 2, and 3 respectively</td>
</tr>
<tr>
<td>EM</td>
<td>Expectation Maximization</td>
</tr>
<tr>
<td>LE1, LE2, LE3</td>
<td>Lab exam 1, 2, and 3 respectively</td>
</tr>
<tr>
<td>LMS</td>
<td>Learning Management System</td>
</tr>
<tr>
<td>Moodle</td>
<td>Modular Object Oriented Developmental Learning Environment</td>
</tr>
<tr>
<td>PT</td>
<td>Progress Tracker plug-in</td>
</tr>
<tr>
<td>PTFQ</td>
<td>Progress Tracker Feedback Questionnaire</td>
</tr>
<tr>
<td>Q1, Q2</td>
<td>Quiz 1 and 2 respectively</td>
</tr>
<tr>
<td>REB</td>
<td>Research Ethics Board</td>
</tr>
<tr>
<td>SCEQ</td>
<td>Student Course Engagement Questionnaire</td>
</tr>
<tr>
<td>SVN</td>
<td>Apache Subversion</td>
</tr>
<tr>
<td>UI</td>
<td>User interface</td>
</tr>
</tbody>
</table>
Table A.2: Conversion to letter and numerical grades from original percentage grades.

<table>
<thead>
<tr>
<th>Letter grade</th>
<th>Percentage range</th>
<th>Median used in numerical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>90-100</td>
<td>95</td>
</tr>
<tr>
<td>A</td>
<td>85-89</td>
<td>87</td>
</tr>
<tr>
<td>A-</td>
<td>80-84</td>
<td>82</td>
</tr>
<tr>
<td>B+</td>
<td>77-79</td>
<td>78</td>
</tr>
<tr>
<td>B</td>
<td>73-76</td>
<td>74.5</td>
</tr>
<tr>
<td>B-</td>
<td>70-72</td>
<td>71</td>
</tr>
<tr>
<td>C+</td>
<td>67-69</td>
<td>68</td>
</tr>
<tr>
<td>C</td>
<td>64-66</td>
<td>65</td>
</tr>
<tr>
<td>C-</td>
<td>60-62</td>
<td>61</td>
</tr>
<tr>
<td>D+</td>
<td>57-59</td>
<td>58</td>
</tr>
<tr>
<td>D</td>
<td>53-56</td>
<td>54.5</td>
</tr>
<tr>
<td>D-</td>
<td>50-52</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>0-49</td>
<td>24.5</td>
</tr>
</tbody>
</table>
Table A.3: Reasoning for questions asked in the PTFQ

<table>
<thead>
<tr>
<th>Question</th>
<th>Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) I regularly checked the graphs in the Progress Tracker.</td>
<td>To see if students actually used the visualizations.</td>
</tr>
<tr>
<td>2) I liked being able to see my progress towards my goal.</td>
<td>To see which of the 3 main aspects of the PT were most liked.</td>
</tr>
<tr>
<td>3) I liked being able to compare my progress on the assignment to the rest of the class.</td>
<td>:</td>
</tr>
<tr>
<td>4) I liked being able to compare my overall grade to the rest of the class.</td>
<td>:</td>
</tr>
<tr>
<td>5) The information provided through the plug-in made me more competitive.</td>
<td>To see if students felt that the visualizations increased competition between students.</td>
</tr>
<tr>
<td>6) The information provided through the plug-in made me more engaged in the assignment.</td>
<td>To see if students felt that the visualizations increased their engagement.</td>
</tr>
<tr>
<td>7) Knowing how the rest of the class was doing helped motivate me.</td>
<td>To see if students felt motivated by knowing the progress of their peers.</td>
</tr>
<tr>
<td>8) Being reminded of my own progress towards my goal helped motivate me.</td>
<td>To see if students felt motivated by seeing their own progress.</td>
</tr>
<tr>
<td>9) I liked the look/aesthetics of the plug-in.</td>
<td>To see if students liked the aesthetics of the plug-in.</td>
</tr>
<tr>
<td>10) I would use this plug-in in all my courses.</td>
<td>To see if students would want the plug-in more permanently integrated into the e-learning system.</td>
</tr>
<tr>
<td>Open-ended questions</td>
<td></td>
</tr>
<tr>
<td>11) What was your favourite aspect of the plug-in?</td>
<td>To find out which aspect students liked best (in case it was not specified in a statement).</td>
</tr>
<tr>
<td>12) Is there anything you would change about the plug-in? If so what?</td>
<td>To allow students to give suggestions for improvement.</td>
</tr>
<tr>
<td>13) Is there anything you did not like about the plug-in? If so, what?</td>
<td>:</td>
</tr>
<tr>
<td>14) Is there a reason you chose not to use the plug-in? If so, what?</td>
<td>To allow students who did not use the plug-in at all to give an explanation why.</td>
</tr>
</tbody>
</table>
Appendix B

Ethics

Table B.1: Data Collection Plan

<table>
<thead>
<tr>
<th>Data</th>
<th>Description</th>
<th>Identity Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online content viewing logs (downloaded from the Moodle website)</td>
<td>The participant’s interactions with Moodle content were monitored throughout the course of the semester using Moodle’s built in logging feature. The original log entries included the course, the time of the interaction, IP address, the student’s full name, the student’s action (view, submit, etc.) and the item or module the student was interacting with (assignment, quiz, lesson etc.)</td>
<td>The student’s IP address and name were removed from the data, replaced by a participant ID number. In order to prevent any possible back-tracing of the log data to an individual student all times were rounded to the nearest hour. Because of the class size and the large number of logged events, this makes connecting a participant back to a student’s data very difficult.</td>
</tr>
<tr>
<td>Assignment submission times (downloaded from the Moodle website)</td>
<td>All assignments for CIS*2500 required a collaboration statement submitted by the student via Moodle at the time of completion for each assignment. The time of this submission was recorded as the submission time for each assignment.</td>
<td>As with the above attributes, any times were rounded to the nearest hour, making it very difficult to link a participants assignment information back to a specific student.</td>
</tr>
</tbody>
</table>

Continued on next page
Table B.1 – Continued from previous page

<table>
<thead>
<tr>
<th>Data</th>
<th>Description</th>
<th>Identity Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment start times (downloaded from the Moodle website)</td>
<td>The time a participant started an assignment was recorded as the first time they viewed the assignment description page on Moodle.</td>
<td>As with the submission times, these times were also rounded to the nearest hour to protect the identification of the participants.</td>
</tr>
<tr>
<td>Grades (provided by the instructor of CIS*2500)</td>
<td>The graded portions of the CIS*2500 class consisted of assignments, lab exams, quizzes, clicker marks, and the final exam. These items were graded on a percentage scale and were recorded within Moodle.</td>
<td>Grades were transformed to the letter grade scale used by the University of Guelph prior to the use in this study (<a href="http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-grds.shtml">http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-grds.shtml</a>). This was to avoid any back-tracing participant identities from exact percentage grades.</td>
</tr>
<tr>
<td>Progress tracker goals (extracted from the Moodle database)</td>
<td>Through the Progress Tracker students were able to set goals for themselves in the course. The goal(s) set by each student were recorded along with the date they were set.</td>
<td>The students' IDs were replaced with participant IDs and times were rounded to the nearest hour.</td>
</tr>
<tr>
<td>Student course engagement questionnaire (SCEQ) (downloaded from the Moodle website)</td>
<td>This questionnaire was administered to all students in CIS*2500 in order to collect self-reported measures of engagement for each student within that class. This questionnaire was voluntary and therefore was not completed by all students or participants.</td>
<td>The SCEQ data was recorded using only the participant IDs.</td>
</tr>
<tr>
<td>Clicker responses (provided by the instructor of CIS*2500)</td>
<td>Students in CIS*2500 were required to bring “clickers” to each class in order to answer multiple choice questions asked by the instructor. The use of their clicker contributes to their participation grade in the class. Questions asked in class were related to the current class material and lessons as well as an aid to identifying topics where the class was struggling.</td>
<td>Clicker data was stored using participant ID numbers. Any student IDs or names were removed.</td>
</tr>
</tbody>
</table>
Appendix C

Data

Table C.1: Attribute descriptions

<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>Description</th>
<th>Post-Cleaning Derivation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>The experimental ID number for each participant. IDs were randomly assigned 3-digit numbers unique to each participant.</td>
<td>Original data.</td>
<td>100 to 999</td>
</tr>
<tr>
<td>First goal</td>
<td>The first steady goal set by the participant using the progress tracker. The goal was recorded if it remained steady for 5 minutes. This was done because some participants tended to rapidly alter their goal, likely in order to view the resulting changes in the graphs.</td>
<td>Converted to letter grade.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Last goal</td>
<td>The last goal set by the participant before the end of the observation period.</td>
<td>Converted to letter grade.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Goal met</td>
<td>Whether or not the participant’s final grade met their first goal.</td>
<td>$grade \geq goal \Rightarrow Y$</td>
<td>Y or N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$grade &lt; goal \Rightarrow N$</td>
<td></td>
</tr>
<tr>
<td>SCEQ</td>
<td>The participant’s score on the Student Course Engagement Questionnaire.</td>
<td>$SCEQScore/104$</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Skills</td>
<td>The participant’s score on questions 1 to 9 of the SCEQ.</td>
<td>$skillsScore/36$</td>
<td>A+ to F</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>Description</th>
<th>Post-Cleaning Derivation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional</td>
<td>The participant’s score on questions 10 to 14 of the SCEQ.</td>
<td>emotionalScore/20</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Participation</td>
<td>The participant’s score on questions 15 to 21 of the SCEQ.</td>
<td>participationScore/28</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Performance</td>
<td>The participant’s score on questions 22 to 24 of the SCEQ.</td>
<td>performanceScore/12</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Engaged</td>
<td>The participant’s self-reported level of engagement in the course (on a 5-point scale).</td>
<td>Original data.</td>
<td>“Not at all” to “extremely”</td>
</tr>
<tr>
<td>Engagement comparison</td>
<td>The participant’s self-reported level of engagement compared to that of their other courses (on a 5-point scale).</td>
<td>Original data.</td>
<td>“A lot less” to “a lot more”</td>
</tr>
<tr>
<td>Priority</td>
<td>Whether the participant preferred to get a good grade or be challenged in the class.</td>
<td>Original data.</td>
<td>“grade” or “challenged”</td>
</tr>
<tr>
<td>Final</td>
<td>The participant’s final grade in the class.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Accumulated</td>
<td>The participant’s accumulated grade before the final exam.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>A1/A2/A3</td>
<td>The participant’s grade on assignments 1, 2, and 3 respectively.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Assign avg</td>
<td>The average of the participant’s three assignment grades.</td>
<td>(A1 + A2 + A3)/3</td>
<td>A+ to F</td>
</tr>
<tr>
<td>LE1/LE2/LE3</td>
<td>The participant’s grade on lab exams 1, 2, and 3 respectively.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Lab avg</td>
<td>The weighted average of the participant’s three lab exam grades (worth 5%, 10%, and 10%).</td>
<td>(LE1 × 0.5 + LE2 + LE3)/2.5</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Q1/Q2</td>
<td>The participant’s grade on quiz 1 and 2 respectively.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Quiz avg</td>
<td>The weighted average of the participant’s two quiz grades (worth 5% and 10%).</td>
<td>(Q1 × 0.5 + Q2)/1.5</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Clicker</td>
<td>The participant’s participation grade in the class, contributed to by their “clicker” activity during lectures.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>Exam</td>
<td>The participant’s grade on the final exam.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>Description</th>
<th>Post-Cleaning Derivation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT views</td>
<td>The total number of times the participant viewed the main progress tracker page throughout the observation period.</td>
<td>Original data.</td>
<td>0-34</td>
</tr>
<tr>
<td>Forum views</td>
<td>The total number of times the participant viewed the forums throughout the observation period.</td>
<td>Original data.</td>
<td>34-3903</td>
</tr>
<tr>
<td>C1/C2/C3 updates</td>
<td>The number of times the participant updated their checklists for assignment 1, 2, and 3 respectively.</td>
<td>Original data.</td>
<td>0 to 69</td>
</tr>
<tr>
<td>C1/C2/C3 complete</td>
<td>The completion percentage of the participant’s checklist for assignment 1, 2, and 3 respectively.</td>
<td>Original data.</td>
<td>A+ to F</td>
</tr>
<tr>
<td>A1/A2/A3 commits</td>
<td>The number of git commits on assignments 1, 2, and 3 respectively.</td>
<td>Original data.</td>
<td>0 to 127</td>
</tr>
<tr>
<td>A1/A2/A3 git submit</td>
<td>The number of hours between when an assignment was due and the participant’s last commit of that assignment to git (negative values represent hours late.)</td>
<td>$assignDue - lastCommit$</td>
<td>-1050h to 160h</td>
</tr>
<tr>
<td>A1/A2/A3 start</td>
<td>The number of hours between when an assignment was posted on Moodle and when a participant first viewed the assignment page (lower numbers mean the participant started earlier.)</td>
<td>$assignView - assignPost$</td>
<td>-15h to 783h</td>
</tr>
<tr>
<td>A1/A2/A3 submit</td>
<td>The number of hours between the assignment due date and when the participant submitted their collaboration statement on Moodle (negative values represent hours late.)</td>
<td>$assignDue - moodleSubmit$</td>
<td>-276h to 152h</td>
</tr>
<tr>
<td>Total access</td>
<td>The total number of times the participant accessed the CIS*2500 Moodle page.</td>
<td>Original data.</td>
<td>310 to 7385</td>
</tr>
<tr>
<td>A1/A2/A3 access</td>
<td>The average number of Moodle accesses per day during the span of each assignment (from when it was posted to the due date).</td>
<td>$daysDuringSpan / accessDuringSpan$</td>
<td>2.5 to 93.5</td>
</tr>
</tbody>
</table>

Continued on next page
Table C.1 – Continued from previous page

<table>
<thead>
<tr>
<th>Attribute(s)</th>
<th>Description</th>
<th>Post-Cleaning Derivation</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1-A2/A2-A3 access</td>
<td>The average number of Moodle accesses per day during the time between assignments (after the previous assignment was due but before a new one had been posted).</td>
<td>$\frac{daysDuringSpan}{accessDuringSpan}$</td>
<td>0 to 78.8</td>
</tr>
<tr>
<td>Exam/Grades access</td>
<td>The average number of Moodle accesses from the last assignment due date until the time of the exam, and from the time of the exam until official grades were released.</td>
<td>$\frac{daysDuringSpan}{accessDuringSpan}$</td>
<td>0 to 92.5</td>
</tr>
<tr>
<td>L1.1 to L12.3</td>
<td>Each of these attributes represents the participant’s lecture attendance. The number after the L is the week number in the semester and the number after the decimal is the lecture number in that week (for instance, L1.1 shows the attendance data for lecture one of week one).</td>
<td>Original data.</td>
<td>Y or N</td>
</tr>
<tr>
<td>Total attended</td>
<td>The total number of lectures attended by the participant (of those with attendance information).</td>
<td>$\sum L1.1+L1.2+...+L12.3$</td>
<td>2 to 25</td>
</tr>
</tbody>
</table>

Table C.2: Data collected

<table>
<thead>
<tr>
<th>Attribute</th>
<th># Responses</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>First goal</td>
<td>58</td>
<td>64.5</td>
<td>(C)</td>
<td>87</td>
<td>(A)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>Last goal</td>
<td>58</td>
<td>51</td>
<td>(D-)</td>
<td>87</td>
<td>(A)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>SCEQ</td>
<td>22</td>
<td>24.5</td>
<td>(F)</td>
<td>68</td>
<td>(C+)</td>
<td>82</td>
<td>(A-)</td>
</tr>
<tr>
<td>Skills</td>
<td>22</td>
<td>24.5</td>
<td>(F)</td>
<td>68</td>
<td>(C+)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>Emotional</td>
<td>22</td>
<td>24.5</td>
<td>(F)</td>
<td>72.75</td>
<td>(B-)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>Participation</td>
<td>22</td>
<td>24.5</td>
<td>(F)</td>
<td>64.5</td>
<td>(C)</td>
<td>74.5</td>
<td>(B)</td>
</tr>
<tr>
<td>Performance</td>
<td>22</td>
<td>24.5</td>
<td>(F)</td>
<td>64.5</td>
<td>(C)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>Final</td>
<td>75</td>
<td>24.5</td>
<td>(F)</td>
<td>78</td>
<td>(B+)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>Accumulated</td>
<td>75</td>
<td>24.5</td>
<td>(F)</td>
<td>78</td>
<td>(B+)</td>
<td>95</td>
<td>(A+)</td>
</tr>
<tr>
<td>A1</td>
<td>75</td>
<td>24.5</td>
<td>(F)</td>
<td>87</td>
<td>(A)</td>
<td>95</td>
<td>(A+)</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Attribute</th>
<th># Responses</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>75</td>
<td>24.5 (F)</td>
<td>87 (A)</td>
<td>95 (A+)</td>
<td>75.33</td>
<td>610.85</td>
<td>24.72</td>
</tr>
<tr>
<td>A3</td>
<td>75</td>
<td>24.5 (F)</td>
<td>61 (C-)</td>
<td>95 (A+)</td>
<td>55.86</td>
<td>776.96</td>
<td>27.87</td>
</tr>
<tr>
<td>Assign avg</td>
<td>75</td>
<td>24.5 (F)</td>
<td>74.5 (B)</td>
<td>95 (A+)</td>
<td>69.47</td>
<td>529.96</td>
<td>21.97</td>
</tr>
<tr>
<td>LE1</td>
<td>75</td>
<td>24.5 (F)</td>
<td>95 (A+)</td>
<td>95 (A+)</td>
<td>85.03</td>
<td>436.03</td>
<td>20.88</td>
</tr>
<tr>
<td>LE2</td>
<td>75</td>
<td>24.5 (F)</td>
<td>95 (A+)</td>
<td>95 (A+)</td>
<td>78.36</td>
<td>599.25</td>
<td>24.48</td>
</tr>
<tr>
<td>LE3</td>
<td>75</td>
<td>24.5 (F)</td>
<td>74.5 (B)</td>
<td>95 (A+)</td>
<td>76.79</td>
<td>223.71</td>
<td>14.96</td>
</tr>
<tr>
<td>Lab avg</td>
<td>75</td>
<td>24.5 (F)</td>
<td>82 (A-)</td>
<td>95 (A+)</td>
<td>78.89</td>
<td>370.07</td>
<td>19.42</td>
</tr>
<tr>
<td>Q1</td>
<td>75</td>
<td>24.5 (F)</td>
<td>24.5 (F)</td>
<td>95 (A+)</td>
<td>42.67</td>
<td>627.90</td>
<td>25.06</td>
</tr>
<tr>
<td>Q2</td>
<td>75</td>
<td>24.5 (F)</td>
<td>74.5 (B)</td>
<td>95 (A+)</td>
<td>70.03</td>
<td>521.98</td>
<td>22.85</td>
</tr>
<tr>
<td>Quiz avg</td>
<td>75</td>
<td>24.5 (F)</td>
<td>64.5 (C)</td>
<td>95 (A+)</td>
<td>59.45</td>
<td>489.27</td>
<td>22.19</td>
</tr>
<tr>
<td>Clicker</td>
<td>75</td>
<td>24.5 (F)</td>
<td>95 (A+)</td>
<td>95 (A+)</td>
<td>91.78</td>
<td>107.02</td>
<td>10.34</td>
</tr>
<tr>
<td>Exam</td>
<td>75</td>
<td>24.5 (F)</td>
<td>74.5 (B)</td>
<td>95 (A+)</td>
<td>73.73</td>
<td>225.16</td>
<td>15.01</td>
</tr>
<tr>
<td>PT views</td>
<td>75</td>
<td>0</td>
<td>4</td>
<td>34</td>
<td>6.83</td>
<td>56.63</td>
<td>7.53</td>
</tr>
<tr>
<td>Forum views</td>
<td>75</td>
<td>34</td>
<td>218</td>
<td>3903</td>
<td>421.40</td>
<td>301316.54</td>
<td>548.92</td>
</tr>
<tr>
<td>C1 updates</td>
<td>75</td>
<td>0</td>
<td>13</td>
<td>39</td>
<td>13.89</td>
<td>61.37</td>
<td>7.83</td>
</tr>
<tr>
<td>C2 updates</td>
<td>75</td>
<td>0</td>
<td>15</td>
<td>69</td>
<td>14.17</td>
<td>168.44</td>
<td>12.98</td>
</tr>
<tr>
<td>C3 updates</td>
<td>75</td>
<td>0</td>
<td>13</td>
<td>66</td>
<td>14.33</td>
<td>138.93</td>
<td>11.79</td>
</tr>
<tr>
<td>C1 complete</td>
<td>74</td>
<td>0 (F)</td>
<td>95 (A+)</td>
<td>95 (A+)</td>
<td>75.83</td>
<td>1027.61</td>
<td>32.06</td>
</tr>
<tr>
<td>C2 complete</td>
<td>71</td>
<td>0 (F)</td>
<td>87 (A)</td>
<td>95 (A+)</td>
<td>64.28</td>
<td>1428.12</td>
<td>37.79</td>
</tr>
<tr>
<td>C3 complete</td>
<td>72</td>
<td>0 (F)</td>
<td>68 (C+)</td>
<td>95 (A+)</td>
<td>61.56</td>
<td>117.49</td>
<td>33.43</td>
</tr>
<tr>
<td>A1 commits</td>
<td>75</td>
<td>0</td>
<td>2</td>
<td>90</td>
<td>4.64</td>
<td>122.94</td>
<td>11.09</td>
</tr>
<tr>
<td>A2 commits</td>
<td>75</td>
<td>0</td>
<td>3</td>
<td>127</td>
<td>6.33</td>
<td>248.55</td>
<td>15.77</td>
</tr>
<tr>
<td>A3 commits</td>
<td>75</td>
<td>0</td>
<td>7</td>
<td>29</td>
<td>7.96</td>
<td>31.61</td>
<td>5.62</td>
</tr>
<tr>
<td>A1 git submit</td>
<td>70</td>
<td>−542</td>
<td>13</td>
<td>161</td>
<td>14.62</td>
<td>5455.11</td>
<td>73.86</td>
</tr>
<tr>
<td>A2 git submit</td>
<td>71</td>
<td>−1050</td>
<td>6</td>
<td>114</td>
<td>−29.29</td>
<td>20392.50</td>
<td>142.80</td>
</tr>
<tr>
<td>A3 git submit</td>
<td>72</td>
<td>−321</td>
<td>4</td>
<td>61</td>
<td>−7.91</td>
<td>4043.17</td>
<td>63.59</td>
</tr>
<tr>
<td>A1 start</td>
<td>75</td>
<td>−15</td>
<td>44</td>
<td>510</td>
<td>111.48</td>
<td>20037.61</td>
<td>140.66</td>
</tr>
<tr>
<td>A2 start</td>
<td>75</td>
<td>0</td>
<td>50</td>
<td>783</td>
<td>258.83</td>
<td>90910.73</td>
<td>299.85</td>
</tr>
<tr>
<td>A3 start</td>
<td>75</td>
<td>0.25</td>
<td>102.25</td>
<td>614.25</td>
<td>212.94</td>
<td>46413.31</td>
<td>215.24</td>
</tr>
<tr>
<td>A1 submit</td>
<td>71</td>
<td>−276</td>
<td>10</td>
<td>152</td>
<td>0.58</td>
<td>2015.93</td>
<td>48.54</td>
</tr>
<tr>
<td>A2 submit</td>
<td>75</td>
<td>−4</td>
<td>12</td>
<td>81</td>
<td>12.83</td>
<td>251.39</td>
<td>15.82</td>
</tr>
<tr>
<td>A3 submit</td>
<td>68</td>
<td>−26</td>
<td>7</td>
<td>133</td>
<td>14.65</td>
<td>705.00</td>
<td>26.36</td>
</tr>
<tr>
<td>Total access</td>
<td>75</td>
<td>310</td>
<td>888</td>
<td>7385</td>
<td>1235.57</td>
<td>1036321.63</td>
<td>1018.00</td>
</tr>
<tr>
<td>A1 access</td>
<td>75</td>
<td>3.8</td>
<td>11.4</td>
<td>93.5</td>
<td>14.99</td>
<td>247.22</td>
<td>12.75</td>
</tr>
<tr>
<td>A1-A2 access</td>
<td>75</td>
<td>0</td>
<td>13.1</td>
<td>61.7</td>
<td>15.32</td>
<td>224.00</td>
<td>13.95</td>
</tr>
<tr>
<td>A2 access</td>
<td>75</td>
<td>2.8</td>
<td>7.8</td>
<td>48.3</td>
<td>9.72</td>
<td>67.94</td>
<td>6.89</td>
</tr>
<tr>
<td>A2-A3 access</td>
<td>75</td>
<td>0</td>
<td>4.5</td>
<td>78.8</td>
<td>9.61</td>
<td>230.30</td>
<td>12.85</td>
</tr>
</tbody>
</table>

Continued on next page
Table C.2 – Continued from previous page

<table>
<thead>
<tr>
<th>Attribute</th>
<th># Responses</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3 access</td>
<td>75</td>
<td>2.5</td>
<td>9.3</td>
<td>88.6</td>
<td>14.06</td>
<td>251.46</td>
<td>13.25</td>
</tr>
<tr>
<td>Exam access</td>
<td>75</td>
<td>0.2</td>
<td>5.6</td>
<td>92.5</td>
<td>12.17</td>
<td>316.19</td>
<td>15.70</td>
</tr>
<tr>
<td>Grades access</td>
<td>75</td>
<td>0</td>
<td>2.2</td>
<td>32.6</td>
<td>4.43</td>
<td>25.60</td>
<td>5.73</td>
</tr>
<tr>
<td>Total attended</td>
<td>75</td>
<td>2</td>
<td>22</td>
<td>25</td>
<td>21.21</td>
<td>13.17</td>
<td>3.63</td>
</tr>
</tbody>
</table>
Table C.1: Histograms of grade data.
Figure C.2: Spearman rank-order correlation between for aggregated data.
References


K. A. Papanikolaou (Eds.), *Proceedings of the 11th International Conference on User Modeling* (pp. 24–32).


clustering for predicting final marks based on student participation in forums. 1–4.


Naps, T., ling, G. R., Anderson, J., Cooper, S., Dann, W., Fleischer, R.,


