CLUSTERING STUDENT INTERACTION DATA USING BLOOM'S TAXONOMY TO FIND PREDICTIVE READING PATTERNS

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Computer Science

University of Saskatchewan

Saskatoon

By

Terry Edward Peckham

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5C9

ABSTRACT

In modern educational technology we have the ability to capture click-stream interaction data from a student as they work on educational problems within an online environment. This provides us with an opportunity to identify student behaviours within the data (captured by the online environment) that are predictive of student success or failure. The constraints that exist within an educational setting provide the ability to associate these student behaviours to specific educational outcomes. This information could be then used to inform environments that support student learning while improving a student's metacognitive skills.

In this dissertation, we describe how reading behaviour clusters were extracted in an experiment in which students were embedded in a learning environment where they read documents and answered questions. We tracked their keystroke level behaviour and then applied clustering techniques to find pedagogically meaningful clusters. The key to finding these clusters were categorizing the questions as to their level in Bloom's educational taxonomy: different behaviour patterns predicted success and failure in answering questions at various levels of Bloom. The clusters found in the first experiment were confirmed through two further experiments that explored variations in the number, type, and length of documents and the kinds of questions asked. In the final experiment, we also went beyond the actual keystrokes and explored how the pauses between keystrokes as a student answers a question can be utilized in the process of determining student success.

This research suggests that it should be possible to diagnose learner behaviour even in "ill-defined" domains like reading. It also suggests that Bloom's taxonomy can be an important (even necessary) input to such diagnosis.

ACKNOWLEDGEMENTS

I would like to acknowledge God who is the giver of all good things. My wife, Jeannette, who gave sacrificially of her time, energy and effort to allow me to pursue this path. I could not have done this without her love and support. My children Jacob, and Joshua who also gave up their dad to allow him the time to reach this goal in his life. I cannot replace the time in all your lives that has been spent working on this project, but I can thank you from the bottom of my heart and acknowledge your sacrifice.

I would like to thank my supervisor Dr. Gord McCalla who spent countless hours talking within me, reading my work, planning, correcting, directing, and revising my work. This work has a lot of you in here as well. Thank you so very much.

My committee: Jim Greer, Richard Schwier, Julita Vassileva, and Ian McQuillan. Thank you for your time and effort in reading this work, providing guidance, and your suggestions on how to best proceed. I truly appreciate your efforts and work.

Finally, I would like to thank my external examiner Dr. Kinshuk. Thank you for volunteering, for your time and for your questions. They were thoughtful and insightful.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

Learning technology is an area that is growing within education and industry as life-long learning becomes more prevalent within our world [1]. People are often required to continue learning throughout their career. The rate of change and the amount of data and information continue to increase at a rapid pace, leading to new technologies, methodologies, and changes within our workplace that keep learning at the forefront. One of the primary methods of learning is in online environments. However, one of the major problems with online learning is that students are often left to figure out concepts and problems on their own without the benefit of having a tutor or instructor help the student or correct their misconceptions. The field of Intelligent Tutoring Systems (ITSs) provides research and insight on how to help mitigate this problem by creating and providing online environments that attempt to provide computer based tutoring systems that can help the students and correct misconceptions. Allied fields such as Advanced Learning Systems (ALS), Artificial Intelligence in Education (AIED), Educational Data Mining (EDM), Learning Analytics (LA), Learning Sciences, Computer Supported Collaborative Learning (CSCSL), etc., also provide a rich source of ideas and techniques.

The digital revolution is also providing opportunities to perform science that could not be explored in the past. In particular, it is possible to record large amounts of student interaction data as students learn. The field of data mining has provided many different algorithms and tools that can be used to search through these large datasets and find pedagogically useful patterns. The constraints of the educational domain provide the possibility to make use of the patterns we find in student interaction data, unlike other settings that are not so constrained. Within an educational domain, we have precise knowledge of a number of factors that can help us to contextualize a student's behaviour:

- which problem the student is currently working on
- the difficulty level of the problem
- what documents, videos, and other sources the student could be consulting

- strategies the student can use to solve the problem they are working on
- and the nature of the answer that the student is trying to achieve

These factors provide us with some constraints that allow us to more readily draw conclusions about the patterns we find in the student interaction data.

1.1 Summary of the Research

Although tutoring systems have made progress in providing students with automated help and in some "well defined" domains are coming close to rivaling human tutors, there is still a lot of work that can be done especially in "ill-defined" domains. One such domain is reading comprehension. In almost any domain, the student requires some method to access and internalize information, and that method is primarily done through reading. Yes, there are videos, animations, sound files, and pictures that present content to the student, but the majority of learning material found in the educational sector is text based, making reading comprehension one of the most critical components in most learning situations.

Our work has demonstrated that we are able to make predictions of student success for various educational problems by watching how the student reads and answers questions about the content presented. In particular we have demonstrated that fine-grained clickstream data as students answer questions can be analyzed to create predictions about successful vs unsuccessful reading strategies. Our work has further shown that different objectives will often require different levels of student comprehension ranging from shallow to deep understanding of the content. These different levels of comprehension will require different types of reading, and these are detectable by the system when looking at the student's interaction data.

Since an educator can ask more than one type of question or assign more than one type of task based on the same document that the student is studying, the task the student is asked to do essentially frames a context for interpreting the student behaviour. What is the effect of these different framing contexts for the same document? Our approach has demonstrated that through analyzing data from students as they read documents and type answers combined with an appropriate categorization of the task in terms of its Bloom level, we are able to successfully predict student letter grades on the questions they are answering. Our successful use of Bloom's

Taxonomy in a technical context is perhaps a role model for its use in other advanced learning technologies.

1.2 Motivation

The three Rs (Reading, wRiting, and aRithmetic) form the basic foundations of education within schools [2]. Each of these individual areas has been the subject of research throughout history dating back as early as AD 401 in the book "The Confessions of Saint Augustine" [3]. The success of a student, often measured by grades, can be traced back to how well a student makes use of their reading and writing skills within academic situations. The field of advanced learning systems (ALS) has built many different tutoring systems that help students gain a deeper understanding of the content they are studying. As previously mentioned, the majority of these ALS are built within "well-defined" domains. Reading and writing by the student are behaviours that are found within these "well-defined" domains and are used to both learn as well as demonstrate their comprehension of the content they are presented with. However, within "ill-defined" domains, the language skills become even more important. The topics found within such domains often require either more detailed explanation in order to demonstrate the concept being presented or present the student with a difficult problem which might have multiple correct responses that cannot be easily defined a priori. In either of these cases, ALSs that can "understand" both the reading and writing behaviours of a student should be able to help a student in a wide variety of both "well-defined" and "ill-defined" domains. By "understand" we mean the system can track the student's behaviour as they read the content and track the student's writing behaviour as they answer questions related to the content, and further the system can find patterns in student behaviour that allow pedagogically useful predictions. In fact, the overall goal of our research is to find these patterns. In particular, we run a series of experiments in which we watch how students read online content, record their interactions as they both read documents and answer questions about what they have read, and then perform data mining on those results with the hope of finding useful patterns in their behavioural data.

Trace methodologies, such as capturing keystroke data, events, eye tracking, etc., have demonstrated that data generated from a student's interaction in an environment can provide information to make cognitive and metacognitive interpretations [4]. This makes sense since

how a student consumes content will have a direct effect on their comprehension of that content. If we know what task the student is currently working on, the difficulty of the task, and the current behaviour of the student as they work on the task, we can make cognitive interpretations [4] [5] [6] [7]. Bloom's Taxonomy [8] of the Cognitive Domain (detailed in Appendix 2), provides a pedagogical framework for determining the cognitively difficult a question/task is. Using this framework we can determine if the student's current cognitive skills, methodology or approach are appropriate for the task that they are currently working on. More specifically, we can map the reading behaviours found within the student interaction data to the corresponding levels of difficulty, as determined by Bloom's taxonomy, of the questions to be answered.

The questions can be categorized as to their Bloom level by using Bloom's Taxonomy Action Verbs [8] [9]. Bloom and Anderson created a list of verbs that direct the way that a question should be answered. These verbs correspond to a level within Bloom's taxonomy. When the action verb is placed at the beginning of the question, it frames the way that the question should be answered [9]. For example, Bloom's lowest level, knowledge, contains the action verb 'list'. Since the goal of tasks at the knowledge level was to remember previously learned information, successfully listing something that the student has previously read would demonstrate that the student has mastered that level of cognitive difficulty for that content.

Data mining techniques are used to discover previously unknown attributes that are informative about a particular set of data [10]. With respect to student interaction data, we are looking for patterns within the data that are predictors as to whether a student will successfully complete some pedagogical task, or not. These predictors may be indicative of a negative behaviour that a student should avoid or positive behaviour that a student should emulate. Many current ALSs make use of data mining techniques to locate specific patterns in student behaviour that can be used to scaffold the student's learning [11].

The patterns that we are most interested in are those that are generalizable beyond some specific pedagogical task. In order to locate that type of pattern, various data mining techniques can be employed. The two primary techniques that are most often used to locate these patterns are association rule mining and clustering techniques. Association rules data mining generates hundreds / thousands of patterns that show up in the data, leaving the subject matter expert to painstakingly examine each of the results to determine if it is a useful pattern or not [11].

Clustering algorithms group the data into related clusters based upon certain attributes that are chosen by the researcher [12]. Since clustering provides a method of grouping similar results together, and we were interested in looking at patterns that appeared across a large number of students, we choose this as an initial data mining approach. We applied clustering to the reading patterns found within the student interaction data in an attempt to see if there were different reading clusters that were predictive of student's grades. The K-Means clustering algorithm is used to locate patterns within student interaction data captured from an online environment as students work on solving educational problems. Currently, we have found patterns that are capable of predicting a student's mark at the letter grade (A,B,C,D,F) level of granularity. These patterns could then be used to inform a student model, advanced learning system, or a teacher so that action can be taken to help a student.

However, the use of the K-Means algorithm is not sufficient to locate predictive patterns by itself. The use of an educational taxonomy such as Bloom's Taxonomy [8] combined with K-Means clustering is necessary to find and locate useful patterns within the student interaction data. The demonstrated use of Bloom's Taxonomy in this manner is one of the contributions of this work to the field of artificial intelligence in education.

Since reading is not a domain-specific task (i.e. it is a skill used in just about any learning domain), it stands to reason that if a student's reading patterns can be used in a predictive way, then we have found a tool that can possibly be generalizable to a large number of domains and as a result have an impact in a great many areas of education. Therefore, the documents that we choose to use for each of the experiments that we run should come from different domains. Furthermore, not only should the domains be different for each of the experiments, but we should also begin to test out different genres of documents (such as poetry, quotations, essays, articles) to see if there are interactions that exist between the genre of a document and how it is read.

One of the important aspects of reading is the context of the person reading the document. For example, if a person is reading a document in order to look up a fact, they will read the document with that context in mind. If a person is reading the same document with the aim to present a treatise on some topic found within the document, the context in which they are reading the document will be different from someone simply looking up a fact. This implies that

the different contexts should result in different reading styles that are detectable by some computerized system. Since the main purpose of reading within an educational situation is to learn something, having students answer questions is important method to determine comprehension from a pedagogical point of view [13] [14] [5]. So one method to change the context in which a person is reading within an educational setting is to create different questions that require different levels of difficulty (according to Bloom's taxonomy) to answer. This will allow us to look for behaviour patterns that predict student success or failure, in terms of grades, for questions at each specific level of Bloom's Taxonomy.

1.3 The Experimental Program

The approach we took to discover the pedagogically useful patterns that exist within student interaction data was to first perform an open-ended exploratory study to see if we could locate these patterns within the student interaction data. We called this first open-ended study Experiment 1. The findings from this first experiment would then be followed up by other experiments (Experiment 2 and Experiment 3) that would both further confirm our findings and at the same time making modifications within each of the experiments to ensure that the findings were robust. Each experiment takes what was learned in the previous experiments and seeks to answer questions that were raised from the previous experiments.

We designed and ran all three experiments using a simple online educational environment as can be seen in Figure 1.1. The top of the interface contains buttons that will allow the user to select the various documents that are included in the study. The text for the documents appears in a constrained textbox just below the buttons with a label indicating the title of the current document selected. On the right hand side are buttons that link to specific questions. When a question button is selected the question and corresponding student answer appear in a box below the article. Finally, a submission button is located at the bottom for when the participant is done the experiment.

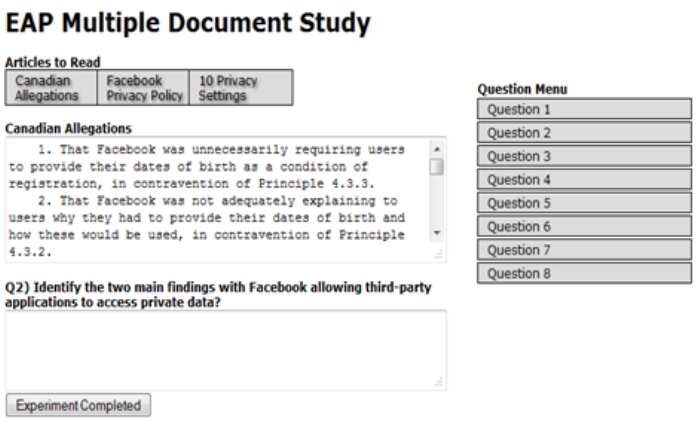


Figure 1.1 A screen shot of our simple educational environment

The purpose of the experiments was to have students work on some typical educational problems that involve reading some content and then answering questions based upon that content. The system recorded as much information as possible within the constraints of an online environment. This included things such as: button clicks, mouse scroll wheel events, the current question the student is working on, the current article they are reading, the position within the article and a timestamp. Figure 1.2 shows a sample of the data that we collected. As can be seen from the PositionWithinTheArticle, Timestamp, and TypeOfEvent fields this particular user is moving quite quickly through the article as they are working on question 1.

ID/PositionWithinTheArticle/Timestamp/CurrentArticle/CurrentQuestion/TypeOfEvent

EX2A700/1/20:57:31:834/A1/Q1/Scroll
EX2A700/3/20:57:31:862/A1/Q1/Scroll

Figure 1.2 Data Recorded from Simple Online Educational Environment

We focused our efforts on what we could infer about the student's reading behaviour by looking at this interaction data. The number of ways that a person can interact with an online document are limited, with the primary one being *reading* the document in full. However, there are other types of reading behaviours such as *scanning* quickly through the text, and a mouse will also allow for *scrolling* behaviour as well as the student moves even more quickly through the document. Both scanning and scrolling can be seen as types of reading differing in the amount of time that is spent taking in information (reading being the slowest and the other two progressively faster – see Chapter 3). As a result, we derived a reading scanning and scrolling

ratio that is comprised of the total amount of time that a student spent working on a particular question. Figure 1.3 provides a sample of those ratios. If the ratios are totalled, you will notice that they do not add up to 100. The portion of time that the student spent typing the answers was not included in this ratio. These ratios will be described in more detail in Chapter 3.

```
EX2B4208
Question: Q1
Reading Ratio: 0.29831353348262296
Scanning Ratio: 0.1502890904720335
Scrolling Ratio: 0.18013436064827115
```

Figure 1.3 Sample of reading ratios

This thesis covers a set of seemingly diverse topics as we seek to integrate topics such as reading comprehension, educational taxonomies, data mining, and ill-defined domains. As a result, the next sections will provide a high level overview of the experiments we performed to frame a context for the material covered in subsequent chapters.

1.3.1 Overview of Experiment 1

The first experiment is an open-ended exploratory study that seeks to determine if it is possible to automatically discover reading comprehension behaviours that exist within student interaction data. Our major research question was: *Is there a computationally feasible method to locate reading comprehension behaviours within student interaction data that are predictive of student success or failure?* We would gauge student success or failure by the student's resulting grade. Although there are other tangible and intangible factors that could be examined, such as increased reading skill, metacognitive problem-solving skills, affect, etc., we decided to focus on grades since a grade is often the summative result of the learning process within educational contexts. A second research question we were seeking to answer was: *Do educational taxonomies such as Bloom's Taxonomy play a role in helping to identify reading comprehension behaviours?*

The first experiment was divided into two separate sections: Experiment 1A and 1B. Since obtaining participants can be problematic, we opted to have two separate scenarios for the participants to work on. This kept experiment 1A more tractable to increase the number of participants. For example, we had a number of participants volunteer for experiment 1A but who

did not participate in the longer experiment 1B. The environment for experiment 1A consisted of one fairly technical document and 4 simple questions for students to answer based upon that document. Experiment 1A was quick and easy to complete. Experiment 1B used the same document from 1A plus two additional documents. The questions asked in 1B were more difficult questions and as a result was a longer experiment to participate in. Those participants that took part in experiment 1B were a subset of those who participated in experiment 1A.

After running the experiment, we performed a modified version of the K-Means clustering algorithm and found some interesting clusters that were predictive of student success in some manner. These clusters formed the starting centroids for our clustering in our other two experiments. Some of our results were as follows:

- The clusters were predictive of the letter grade that a student received for an individual question.
- Statistically interesting clusters could only be found if the data also included the level of difficulty of the question as indicated by Bloom's Taxonomy.
- Students who did not adapt their reading behaviour to the Bloom level of a task performed poorly compared to those students who did modify their reading behaviour for that task.
- Heavy Reading (as defined in Chapter 4) was not always the best method to achieve good grades.

These results were extremely interesting and opened the door to many other avenues to pursue.

1.3.2 Overview of Experiment 2

In order to be sure of the reliability of our results, and to ensure that we had not over-fitted the results, we needed to confirm the results of the first experiment with another experiment. The major research question for this experiment was: *Can we discover reading comprehension behaviours within student interaction data that are predictive of student success or failure within a different domain and with different types of students?* Further, we wanted to see whether it might be possible to improve on Bloom's Taxonomy as a measure of the difficulty of a given question. A second research question was: *Can we gain a better grade predictability by making use of a different educational taxonomy that has more levels of cognitive difficulty compared to Bloom's Taxonomy?* In particular, since Bloom's Taxonomy contains six levels of cognitive

difficulty, and using this we could predict letter grades, *would a taxonomy such as Marzano's that has 14 levels of cognitive difficulty be a more accurate predictor of grades?*

The second experiment was designed based upon experiment 1B but used a different set of participants than the first experiment. Again, this experiment was divided into two components called Experiment 2A and 2B. Experiment 2A contained multiple documents, but on a different topic than experiment 1. The questions students were required to answer in experiment 2A were at the lower levels of Bloom's and Marzano's taxonomies. Experiment 2B contained the same documents as 2A but asked questions at the higher levels of Bloom's and Marzano's taxonomies.

In order to address the concern about possible over-fitting, rather than use K-Means to locate our important clusters, we used the 4 hard-coded centroids that we found in experiment 1. This produced the same statistically significant results we found in experiment 1 suggesting that the patterns we found in experiment 1 were real and our results in that experiment were not over-fitted.

Since the number of participants in both experiments was low, we combined the data from both experiment 1 and 2 when we examined if Marzano's Taxonomy provided better prediction. All the questions from both experiments were recasted into Marzano's Taxonomy, and the clustering algorithm was run again, using the centroid found in experiment 1. Even though Marzano's Taxonomy has fourteen levels compared to Bloom's six levels, the use of Marzano's Taxonomy did not improve the predictability of our clusters. The Marzano clusters allowed the same predictability as Bloom clusters had.

1.3.3 Overview of Experiment 3

After experiment 2, we decided not to pursue Marzano's Taxonomy as an alternative to Bloom's Taxonomy. In the third experiment, we wanted to perform a more in-depth look at students' reading behaviour at the higher levels of Bloom's Taxonomy. The first research question that we wanted to look at was: *How well do the results from the first two experiments hold up when subjected to a difficult ill-defined domain question?* With this in mind, we designed an experiment that included a high level Bloom question that required a much longer answer than we had previously requested from participants. The answer to this question needed to be at least one and a half pages in length and required the student to make citations from the

documents read. The question was from an ill-defined domain (it was about cultural identity) and had no real right or wrong answer, although there were, of course, better or worse answers as expressed in a grading rubric. This type of question would test the abilities of the clusters to detect various grades for questions at higher levels of Bloom's Taxonomy. In addition to the one long answer question, three lower level Bloom short answer questions were also posed. A second research question that we had was: *Does the amount of time that a student has to answer the questions affect the behaviours they use as they both read and answer the questions?*

As in the previous experiments 2 the experiment was run in two parts: Experiment 3A and 3B. Experiment 3A was an experiment that was done like experiment 1B and 2B but for a longer period of contiguous time (three hours). Experiment 3B was the same as 3A in all respects except for the time frame. Experiment 3B was run over two weeks during which the students could work at their own pace and in their own time. This was done to see the effects that the experimental time frame had on how students interacted with the documents.

The results of this experiment confirmed the results of the first two experiments with respect to grade predictions for the lower level Bloom questions. The time differences between Experiment 3A and 3B demonstrated that different strategies were employed by those students who performed the two week session compared to the three hour group. We also found that our clusters were not very predictive of grades for the high level Bloom question. However, when we performed clustering on the keystroke data recorded from the students, we did find that it was predictive of grades. More specifically the pause between the keystrokes for different clusters was predictive (75% of the time) of the letter grade. The remaining 25% of the time required a combination of Heavy Reading and the pauses between keystrokes to predict the letter grade. This last finding was surprising and may open an entirely new line of enquiry: *what is happening during these pauses in typing?* This has been left for future research.

1.4 Contributions

There are several main contributions of this research:

1. Including Bloom's Taxonomy as an integral part of the data mining process can be used to locate pedagogically useful patterns within student interaction data.

2. Reading, Scanning, and Scrolling metrics are useful metrics for reading comprehension within the context of an online environment.
3. Specific diagnosis techniques based on K-Means clustering to have been developed that are useful in the reading comprehension domain.
4. We have found four distinct clusters of behaviour that are predictive of grades based on the Reading, Scanning and Scrolling metrics and the cognitive difficulty of the question to be answered.
5. The timing between the keystrokes when students are typing their answers seems to have some predictive capabilities with respect to a student's grade for a particular level of cognitive difficulty.
6. The computational expense of data mining is not necessary for the locating of known patterns. The centroids used for Experiment 2 and 3 relegate the clustering algorithm to a single pass rather than multiple passes found in normal clustering algorithms. This allows us to more rapidly compute results rather than having to resort to data mining.
7. We have shed light into the possibility of using these discoveries across many domains, especially ill-defined domains, since reading is a core skill necessary in so many domains.

1.5 Overview of Dissertation

The rest of the thesis is organized as follows. In order to better contextualize our work, we provide a survey of the reading comprehension field in Chapter 2. Chapter 3 provides a discussion of our experimental design, including the steps of preprocessing the data and a discussion of granularity issues. Chapter 4 describes the first experiment and the results of that experiment. Chapter 5 describes the second experiment, providing confirmation for our first experiment plus looking at new educational frameworks. Chapter 6 describes the third experiment and the results of that experiment along with some discussion of using additional behavioural data, most notably pauses between keystrokes when students are answering questions. Chapter 7 draws some conclusions about the work, speculates on what else could be discovered from the data we have recorded, and outlines some of the broader implications of this research for advanced learning technology, especially for advanced learning systems that have the goal of supporting learning in ill-defined domains.

CHAPTER 2
BACKGROUND AND RELATED WORK

Reading comprehension is an ill-defined field. There is no unified theory of reading comprehension that can be used as the framework upon which we can build an advanced learning system. Since our system is measuring reading comprehension in some manner, does it line up with any of the current theories of reading comprehension? This section will be a brief overview of the broad theories of reading comprehension with an aim to provide some background on reading comprehension theory. It will be followed by a brief history of ITS, and in particular, discuss how natural language processing is used within many ITS frameworks as well as how reading comprehension is used within ITSs. Finally, we will discuss ill-defined domains and constraint based tutoring systems as it relates to our research efforts.

2.1 Broad Theories of Reading Comprehension

Reading comprehension has been defined in many different ways, yet most models suggest that learners engage in three processes:

i. Decoding or recoding
ii. Comprehension
iii. Response

Decoding or recoding can be seen as a process of converting printed language to spoken language regardless of whether it is through overt vocalization or covertly as an inner language. Comprehension involves deriving meaning or mentally modelling the text and is usually viewed at many levels such as literal, inferential or interpretive. Response involves comprehension at the interpretive level with the addition of affect, appreciation, and/or application. This could occur during or after reading [15].

The following sections will briefly discuss the major theories of reading comprehension. The discussion will start with some of the earlier, more general theories and move to the more recent theories. It is these more recent theories that start providing the necessary framework

from which ALSs can start developing modules that incorporate reading comprehension components that may allow expansion into other domains.

2.1.1 Social Cognitive Theory

Albert Bandura proposed Social Cognitive Theory, used in education, psychology, and communication [16]. This theory states that an individual's knowledge is not solely learned by oneself. Rather it is learned within the context of social interactions, experiences and media influences [16]. From an educational perspective, learning occurs through the observation of teachers, parents and peers who act as models for the learner. This model is constructivist in its view of reading comprehension in that the reader, the text, the teacher and the classroom community are all involved in the construction of meaning [17].

Although there are aspects of constructivism in ITSs, the reading comprehension tutors are just now starting to employ constructivist techniques (iSTART-ME) [18]. For the most part, the tutors and the students engage in one-on-one types of activities.

2.1.2 Schema Theory

Schemata can be defined as interacting knowledge structures that shape our expectations for reality [19]. There are two major types of schemata: content and formal schemata. Content schemata refer to the background knowledge that a person possesses about a topic found in the text. Formal schemata refer to the semantic structures that exist in the different types of texts. From a procedural perspective, there is a two-way process between the text and the background knowledge. There is a cognitive element of the schema that connects the incoming information with the reader's prior knowledge about that information. The reading process also involves the processing of the formal schema in terms of genre, topic, and other structures that allow readers to comprehend the text. These structures allow the reader to make predictions on future text and to see if the current structure they are reading matches an existing schema. If it does not, then the reader now must decide if they are going to add it to their schemata or reject it [19] [20]. However, the more life experience the reader has, the more structures they have added to their schemata and therefore, the more prepared they are to make connections within their reading.

Aspects of this theory can be seen inside of the reading comprehension tutors. Minsky implemented a similar frame theory in 1975 and used this as a basic knowledge representation

scheme for an intelligent system [21]. Further refinements of his work were foundational to the development of object oriented systems, agents, etc.

2.1.3 Transactional Theory of Reading and Writing

The transactional model of reading and writing are dynamic in nature and include both aesthetic and cognitive aspects of reading and writing [22]. Every act of reading is called a transaction between the reader and the text within a particular context. The meaning of a passage of text does not reside specifically within the text nor within the reader, but resides within the transaction between the reader and text. Therefore, text without a reader is simply a set of marks on a page or screen capable of being interpreted as written language. However, when a reader transacts with the text, meaning occurs [22]. This same premise forms the foundation of situated cognition with a few more additions [23].

Situated cognition has made in-roads into the AIED field [24] [25] [26]. The context in which content is presented has a direct impact on how that content is encoded by the learner. In ill-defined domains, the context becomes especially important since the same content can have multiple meanings and what the learner needs to encode is therefore dependent on the context. This has direct bearing on our research, since we are looking at how a single document can be used in different contexts and the strategies employed by the learner need to be different depending on the context.

2.1.4 LaBerge and Sammuel's Theory

LaBerge and Sammuel's theory [27] espouses the concept of "automaticity" which is used as a metaphor to explain that some acts are performed by an individual that are beneath the level of conscious awareness [28]. During the reading process the reader learns to encode starting with a letter by letter perception and eventually moving to phonemic parts of the sub-word to finally comprehending the word as a whole [28]. When the reader is able to decode a whole word, the comprehension process is considered to be automatic.

Within the ITS field, there is the ability to capture the keystroke / mouse click data of a student interacting with a system. This keystroke data can then be analyzed to determine the learner's behaviours using the fine-grained data. Our research is based upon taking this fine-grained data and data mining it in order to determine the learner behaviour [29] [30]. However,

there is a limit to the number of patterns that can be gleaned from just fine-grained student interaction data. As a result, some abstraction of this data is required for more meaningful patterns to be discerned in a student's interaction behaviour.

2.1.5 Construction Integration Theory

The Construction-Integration (CI) theory proposed by Kintsch [31] [32] has the goal of being a comprehensive theory for reading comprehension. This theory assumes that there are three main forms of knowledge representation: verbatim, propositional, and situational. Verbatim information is that which is stored in the surface structure, for example, specific words in a sentence or the syntactic structures that exist in the sentence or phrase. Propositional knowledge is abstract in nature and is formed when abstract propositional-schemata are instantiated with surface structure information [15]. These abstract propositions are then merged into a larger propositional text base made up of micro and macro propositions, capturing the situational context. This forms the basis of a multiple coding theory that assumes that verbal language and mental imagery are both input and output constrained by the propositions developed around the base text.

It would be interesting from an advanced learning system perspective to have a mental model for each user model within the ALS to model the three forms of knowledge representation that may be occurring as a student reads. CI theory does provide a framework from which an intelligent tutoring system could begin to tackle the task of reading comprehension.

2.1.6 Dual Coding Theory

Dual Coding Theory (DCT) [33] was developed to account for verbal and non-verbal effects on memory. Rather than being a theory dedicated to reading comprehension it is a theory of the mind that has been applied to a number of domains, including reading comprehension [34]. This theory proposes that mental representations that can be empirically observed and tested in various verifiable ways are scientifically valid while abstract mental mechanisms such as propositions, schemata, or other such mechanisms remain non-valid [15]. All encoding of mental models can be derived from some sensory input and are classified as either verbal or non-verbal (dual codes) that are either linguistic or non-linguistic in their nature [33]. For example, in the auditory and articulatory modalities phonemes, word pronunciations and stress intonations can be classified while in the visual modality, letters, Braille, and punctuation marks can be

classified. Furthermore, there is simultaneous additional sensory input such as sounds, smells, tastes that can accompany the aforementioned representations. The verbal representations tend to be sequentially processed whereas the non-verbal representations tend to be simultaneously processed [15].

DCT relies heavily on behavioural data rather than computer simulations that seek to approximate some behaviour [15]. Within AIED we are also very concerned about the analysis of behavioural data as it applies to student behaviour within some intelligent tutor. At the same time, this theory also seems to be in conflict with many AIED approaches that model student behaviour.

2.2 ITS From a Reading Comprehension and Natural Language Processing Perspective

One of the major goals of an ALS is to foster deeper learning by the students who interact with the ITS. Within the framework of reading comprehension and natural language processing, this has taken a number of forms. These forms include: combining expert systems with natural language processing (NLP), maintaining student interest as they perform reading comprehension tasks, training students in multiple reading strategies, as well as oral and auditory applications of reading comprehension. The following sections provide a brief history of ITSs followed by discussions on the various forms ITSs have taken as they have worked towards fostering deeper learning within reading comprehension and NLP problems.

2.2.1 ITS: The Early Days

In the 1960s several generative Computer Assisted Instructional (CAI) systems were developed [35]. Basically, these systems operated as automated flash cards designed to present the student with a problem, record the student's response, and then calculate the student's performance on that task. These systems did not explicitly address the issue of how people learn, but used a behaviourist / transmission model of teaching and learning [35]. These systems assumed that if a learner was presented with information to learn, the learner would absorb it.

However, by the late 1960s and early 1970s, several researchers moved beyond simply presenting problems to learners and then collecting their responses. Researchers were now beginning to consider the student factor in the overall instructional system [36]. A good example of this type of system is SCHOLAR [37]. These types of systems altered the presentation of

material based on the history of the student's responses. These systems were the first to model students with the main thrust of attempting to diagnose student skills in some sense. The systems were relatively simple by today's standards, yet by constraining themselves to the development of skills and recall; they were effective. The implicit learning theory used by these systems assumed that the students needed to learn the basic skills and facts prior to learning higher order synthesis skills [38].

Throughout the 1960s, 70s and 80s there was great hope for rapid progress within the field of AI. These were spurred on by advances in computational power. However, many of the problems that AI was considering ended up being much more intractable than the challenges of building faster computers. Throughout this period of time the goal of "thinking" computers always seemed to be just 10 years away [39].

Around this time, the field of educational psychology began questioning the assumptions of behaviourism. Piaget's theories of learning and constructivism began to take hold. Chomsky, Newell and others introduced the ideas of symbolic processing that dovetailed with the AI community's interest in linguistics and natural language processing (NLP) [40].

In the introduction to their classic collection of papers about 1970s-era ITSs, Sleeman and Brown [38] set out to describe the differences between CAI and ITSs. They classified existing ITSs as: computer-based problem-solving monitors, coaches, laboratory instructors, and consultants. Within ITSs, there was an explicit move towards the representation of student knowledge. It is here that the term "student model" [41] was first used to describe the representation of the student by the tutoring system. An example early system that made use of such models is DEBUGGY [42] [43]. This system analyzed the problem space represented by a student's answer to determine which bug or set of bugs best accounted for an error during a mathematical subtraction problem.

John Anderson proposed his theory of cognition called the Adaptive Control of Thought (ACT) theory in 1983 [44]. Although this theory was developed as a cognitive theory, Anderson believed that it was rigorous enough to have its principles implemented inside a tutor. Two good examples of such tutors are the Geometry Tutor [45] and LISP Tutor [46]. Corbett and Anderson found that students who used their LISP tutoring system mastered the content

significantly faster than those students who worked alone, but not as fast as students who worked with human tutors [46].

By the mid-1980s an emerging topic of interest in AI was expert systems that worked best in well-defined domains. At this same time, the ITS field was moving out of the laboratories and into the classroom. Rosenberg made note of two major flaws within ITSs that came to light as the research moved into the classroom: The systems were not grounded in a substantiated model of learning. The evaluation ITSs was often based upon severely flawed tests [47].

In the five years following the Sleeman and Brown review, the field of ITS continued to grow and evolve. Wenger, in 1987, called for a move towards a "cognitive oriented form of software engineering" [48] where cognition becomes the central focus rather than the computational models of the domain and pedagogy.

The field of ITS continued to incorporate a wide range of AI techniques. In fact, the applications spread beyond "tutoring" to support learning in many ways, and the field of ITS became part of a much broader field called Artificial Intelligence in Education (AIED). While many techniques are used within this field, one of the defining aspects of most AIED systems (especially ITSs) is the use of individualization [49] [50] [51] [52]. Individualization allows the system to adapt to the learner (through the use of a "learner model," aka student model), one of the keys to stimulating successful learning.

2.2.2 Conversational Dialogues and the Interpretation of Natural Language Responses

There are several examples of intelligent tutors that incorporate dialogue into their pedagogy [53] [54]. The main idea behind this type of tutor is to obtain answers from the student to help demonstrate their proficiency in the content the tutor is teaching. These are often done through textual dialogue; however, there are instances where spoken dialogue is used as well [53] [55]. The examination of a textual response from the student is followed by entering into a dialogue that fosters deeper comprehension of the current task the student is working on. In these cases, the objective of the tutoring system is not to teach or perform reading comprehension but rather to make use of natural language as a means to gauge which content the system should display next based upon the student's response to the tutor.

AutoTutor is an ITS that capitalizes on conversational dialogue between the student and the tutor and is a great example of a tutor that uses conversational dialogues [56]. The main goal of AutoTutor is to support explanation-centered learning through dialogue. The tutor makes use of an animated avatar that interacts with the student. The tutor engages in text based dialogue with the student, and animations, such as a nodding head or a pointing hand, help to reinforce the message. The tutor provides dialogue cues and questions to guide the student through the pre-scripted material that the student is required to learn.

In addition to being able to enter into dialogue with the student, AutoTutor also maintains some ability to answer questions that are posed by the student. In most human based tutoring situations, the majority of the conversation relies on the tutor to initiate the conversation and probing the learner to determine what they know and don't know [57]. However, there are times when the student starts the conversation by a question and AutoTutor is set up to handle some questions. The tutor has six different categories into which it bins questions asked by the student. Each bin has a predetermined routine to follow that both answers a student's question and then continues on with the dialogue. For example, the tutor can answer a question by looking up a matching definition for some keyword found in the student's question. The tutor maintains a glossary of keywords and their corresponding definitions that are related to the content being covered by the tutor. The tutor then presents the definition for the keyword it has determined the student is requesting clarification on. This is helpful in that it can allow a student to better understand the content being presented when they know the definition of a word they may not have previously encountered.

Although this tutor is not focused on reading comprehension, it does make use of dialog interaction and a formal analysis of the student's answers in order to determine if a student has correctly answered a question to see if learning has occurred. This could be considered to be a shallow word-based analysis of student text [58], and could be a useful approach for work that is focused on reading comprehension.

2.2.3 Combining Expert Systems with Natural Language Processing

The process of posing properly constructed questions to a student is known to increase the depth of knowledge that the student retains about some concept [59]. The Atlas tutoring system is a rare example of a system that works in conjunction with other pre-existing tutoring

systems to allow natural language discourse between the tutoring systems and the student. This type of tool allows for the analysis of a student's natural language response to a question and then picks the best content to display to the student based on their response.

Atlas started out as a supplementary tutoring system that worked with other tutoring systems interactively such as Andes, Why2, and IT-SPOKE [56]. The Atlas system made use of knowledge construction dialogues (KCDs) that encouraged students to infer or construct the target knowledge rather than have it given to them. This encouraged much deeper kinds of learning by connecting principles, relating them to common sense knowledge and then getting the students to talk about those connections and principles. Like Andes, Atlas enters into a dialogue with the student. In order to be able to deal with the student's response, the tutor makes use of CARMEL [60] to translate the expected student responses into semantic structures. This allows the tutor to jump ahead through the material if the student is demonstrating more advanced knowledge than where the tutor is currently in the script.

2.2.4 Reading Comprehension and Increased Vocabulary While Maintaining Student Interest

Learning how to read effectively and improve vocabulary is a time consuming task. The best way to improve your reading skill is to read more. The amount of time that is required by a tutoring system to assess a student's improvement in reading comprehension and vocabulary is long. As a result, students often get bored by working within a tutoring system before they have completed the content. To combat this, several tutoring systems have been developed to maintain student motivation as they work through the material [53] [18].

The REAP tutor is one of the first motivational reading comprehension tutoring systems developed [61] [62] [63]. It is a web-based system that helps students learn English as an Alternative Language (EAL) to improve reading comprehension and their vocabulary. The tutor provides reader-specific passages that are based on a user model. The student model includes reading level, topics of interest, and vocabulary goals. Given that the student model contains a list of vocabulary that the user already knows, and what their interests are, it is possible for REAP to search its database for documents that fit their interest, contain vocabulary they already know, plus new vocabulary that they do not know. This gives the student a chance to reinforce what they already know plus grow their vocabulary with new words [61]. The tutor also makes use of the Cepstral Text-To-Speech system as well as the Cambridge Advanced Learner's

Dictionary installed to help students both verbalize the word as well as learn the meaning of new vocabulary words [63]. Following each reading assignment, an assessment is performed to ensure that the student understands the new content and to update the student model and to move on to the next level.

Retaining student interest and attention during long reading comprehension tasks is difficult. It takes time and repetition in order to strengthen a student's skill sets. REAP is one of the first tools to take on this difficult task of maintaining student interest while scaffolding learning. Providing content that a student is interested in learning or reading about in a manner that facilitates learning and vocabulary is a great way to motivate the learner. We do not see much more development in maintaining student interest within other reading comprehension tutors until the iSTART-ME game is developed in 2009 [18] and Protutor in 2010 [53]. This will be discussed further in sections 2.2.5 and 2.2.6.

2.2.5 Training Students in Multiple Reading Strategies to Improve Comprehension

Some tutoring systems are aimed specifically at the task of improving reading comprehension. These tutors are often also categorized as being in the Computer Assisted Language Learning (CALL) field, such as iSTART and Project Listen for example. These tutoring systems use content specifically aimed at improving reading comprehension at multiple levels.

iSTART is an example of a web-based reading strategy tool that makes use of animated agents to model, discuss and provide feedback about the reading strategies used in science texts. iSTART uses the Self-Explanation Reading Training (SERT) intervention [64] [65]. This is the process of explaining the meaning of the text while the student is reading the text through the use of seven different strategies [61]. There are two major reasons for the use of SERT within iSTART. First, students who self-explain the text are more likely to have developed a deeper understanding of the text and more complete mental models of the content [66] [67]. Second, reading strategies promote successful comprehension [68] [69]. The integration of self-explanation with reading strategies builds the necessary toolset for students to understand both difficult and unfamiliar concepts and texts.

iSTART makes use of only five of SERT's seven reading strategies: comprehension monitoring, paraphrasing, making bridging inferences, prediction and elaboration [70]. Comprehension monitoring is the task of having the reader notice when they don't understand the current content, which then triggers the use of other active reading strategies. Paraphrasing is the process of transforming the surface structure of the text into something that is more familiar to the student. Ideally, the student should move beyond simple transformation of the content towards mapping links between the current content being read and other prior knowledge or content. Bridging inferences is the process of linking the current sentence being studied to material that was previously covered. These bridges allow the reader to develop a global perspective of the content being learned. The use of prediction to anticipate future topics in the text helps in the comprehension of the text. It does not matter if the student guesses what will come or if the student reminds himself or herself to actively seek out what is coming next, as both result in better comprehension. Lastly, elaboration associates the current sentence to prior knowledge so as to result in a coherent and stable representation of the content [71] [32].

iSTART-ME is the latest phase of the iSTART tutor [72] [73]. The iSTART-ME tutor is a game based tutor that tries to make use of a gaming environment to get the students to become more actively engaged with the tutor. This is accomplished through interaction using 10 to 20 minute mini-games. The mini-games try to improve a student's identification of reading strategies, generation of self-explanations, vocabulary and meta-comprehension awareness [73].

From the main game menu the students can select options to interact with new texts, earn points, advance through levels, purchase rewards, personalize characters and play mini-games. These features have been shown to increase motivation and learning [73].

The iSTART-ME component, in addition to supporting long-term one-on-one interaction with the tutor, allows for a peer-to-peer component. This is an interesting component that takes aim at some of the constructivist theories of student interaction. The peer interaction helps to further maintain motivation of the students through competition and cooperation.

2.2.6 Oral and Auditory Applications of Reading Comprehension

Some tutors make use of voice recognition software [53] to listen to students and determine if they have pronounced words correctly. Similarly, there are tutors that make use of

text-to-speech systems to read text aloud to students. There are also some tutors that make use of both components to better scaffold reading comprehension.

Project Listen is a guided reading tutor that provides students with a chance to practice vocabulary and comprehension within a defined context [74]. The primary audience for Project Listen is elementary school children and individuals for whom English is an alternative language. The tutor works by selecting a document for the student to read, watching them read the document (by making use of voice recognition software to listen to the student read the material), correcting pronunciation when necessary and providing any necessary definitions when requested [55]. Once the student has read the document they are given a chance to choose another document to read. Poor readers will often choose to reread a document [75] [76]. This helps to reinforce word fluency but does not always help with comprehension. The third document that is read is chosen by the tutor and is a new document to the student that will try to expose the student to new vocabulary [77]. This is done by drawing on Vygotsky's idea of the zone of proximal development [78]. The tutor will dynamically update the tutor's estimate of the student's current reading level based on the vocabulary of the last document the student read and then choose a new document that is somewhat harder than the last document, in terms of the vocabulary found in the document [76]. The tutor also scaffolds comprehension by reading aloud more complex sentences and then asking questions to ensure comprehension [59]. This tutor has been shown to help students in their word recognition, fluency (speed and accuracy of oral reading), comprehension, and spelling [54].

2.2.7 Reading Comprehension Aimed at Fostering Deeper Learning

One of the goals in teaching should be to foster deep learning through the use of questions requiring deep reasoning. Research has demonstrated that the majority of students in the classroom environment don't ask many questions [79]. Furthermore, the questions that are asked tend to be of lower quality. Ironically, during evaluation most teachers pose shallow rather than deep questions since they are easier to mark [80].

An example of a tutoring system that tries to support deep reasoning is Point and Query [81]. The goal of this system is to provide good examples of questions that require deep reasoning as students explore some content. Point and Query is a web based system that allows users to point at a picture and click. Following the click a series of questions populates the

screen allowing the user to select questions about the picture they clicked on to learn more information about it. This will allow the students to actively engage in the topic and learn how to generate good questions that are required to fill in missing information that a student could have about a topic. However, subsequent research demonstrated that it is not sufficient to simply expose students to examples of deeper questions. When left to their own devices, students quickly resort to shallow questions unless they are presented with a deeper goal or task [81].

2.3 Ill-Defined Domains

Ill-defined domains and tasks are often described as being very complex [82]. Some researchers state that ill-defined tasks are those that have no correct solution. Instead, they consist of a family of solutions that can all be deemed correct [82]. Model tracing [83] is the dominant form cognitive tutoring paradigm employed within the well-defined domains of the ITS field. However, model tracing suffers from design issues when applied to ill-defined domains [84]. Since each of the concepts and misconceptions that are used to create the trace through the material are made by experts, the ITS cannot handle any conceptions or misconceptions that are not within its knowledge base [84].

Stellan Ohlsson proposed constraint-based modeling (CBM), in 1992, as a method to overcome some of the problems associated with model tracing [85]. Rather than trying to model every type of interaction, which for ill-defined domains was extremely hard, CBM would create a system that would check to see if the answers violated any constraints [84]. If no constraints were violated in the creation of the solution, then the solution was judged to be correct, even if the solution was unorthodox. The first major constraint-based ITS was the SQL-Tutor and over time it was followed by many different types of tutoring systems [85] [84]. Constraint-based tutoring systems also adapted many of the interesting attributes of their model-tracing brethren [70] [86]. For example, CAPIT is a constraint-based tutoring system that teaches punctuation and capitalization skills to elementary school children [87]. It makes use of a Bayesian network to help predict the best next problem to display to the student as well as feedback functions to ensure that each student gets the most relevant feedback to their current situation [87]. It should also be noted that natural language processing and tutor based dialoguing systems are also found within CBM tutors [84]. Interestingly, both the model-tracing and CBM-based tutoring systems

make use of NLP in the decision making process through the use of adaptive tutorial dialogues to choose the next best response to the student's input [84]. Although constraint based modelling overcomes some of the problems found within the model tracing architectures [84], constraint based modelling also has shortcomings in that, for some domains, there is no easy way to create constraints that allow for modeling of the concept. For example, an assignment allowing students to argue for either side of a specific topic such as portraying cultural diversity in a positive or negative light depending on personal perspective.

2.4 Summary of ITS Field

My work, as outlined in this dissertation, does not neatly fall into any one particular category within the field of ITS. Learning and comprehension behaviours find their roots in educational cognitive theory. My work relies heavily on this premise as we incorporate Bloom's Taxonomy of Cognitive Difficulty [9] [8] as an integral part of the process required to predict student success or failure from student interaction data.

Conversational dialogues such as AutoTutor [57] seek to probe what a learner knows through a questioning process. This questioning process continues with the Socratic like methods of Atlas, Andes, and IT-Spoke [79] [56] that look to foster deeper understanding of the content. The use of questions to gauge the comprehension of a student has a long history. Our work continues this with the asking of questions that are framed within the context of an educational taxonomy. Graesser and others over the years have used questions to move students towards a deeper understanding of the content [79]. Our aim is slightly different. We seek to understand and quantify the behaviours that exist in student interaction data that demonstrate that both deep and shallow learning has occurred. If we can accurately and automatically identify when comprehension has occurred, then the door opens for other ITSs to leverage this knowledge to help the student in the learning process.

Reading is one of the primary ways that students learn new content. Reading comprehension with the ITS field has seen many different forms over the years. These range from learning vocabulary [88] to teaching pronunciation [55]. Tutoring systems such as I-START [89] seek to teach new reading comprehension methods to provide students with new reading techniques. Our work does not seek to teach reading comprehension directly; rather we

are seeking to find patterns within the data that demonstrate that reading comprehension has occurred. Reading is one of the primary forms of relaying concepts within ill-defined domains.

Ohlsson's idea of constraint based modelling is one of the primary methods used by ITSs designed for ill-defined domain [85]. Our approach is not to constrain an environment but rather to observe student behaviour to determine if they have learned. Our approach can be summed up as doing deep educational diagnosis of reading strategies based on tracking fine-grained behaviour of readers as they read. In the next chapter, we will turn to our own work and examine some of the basic assumptions of our approach.

CHAPTER 3
EXPERIMENTAL DISCUSSION, DATA PREPROCESSING AND GRANULARITY DISCUSSION

Prior to discussing each of the experiments in detail, a discussion about the overall methodology used in the experiments, along with several issues surrounding the data and granularity issues found in these experiments, is required. In order to produce interesting results from the data that is collected, preprocessing of the data is one of the necessary steps. This will provide a framework that will allow for a more coherent interpretation of the results of the experiments conducted. Further descriptions of the interface, data mining techniques and experiments will be discussed in subsequent chapters 4, 5, and 6, as each experiment is discussed in detail.

3.1 Experimental Interface

The interface described here and shown in Figure 3.1 was used in all the experiments. For each experiment, the only variance in the interface was the number of questions and the number of documents that were presented to the student. The design was simple and minimalistic to decrease the number of possible behaviours that could be tracked compared to a more complicated interface. To aid in determining what part of the document was currently being read, a small scrollable text box was provided for the user to "look through" as s/he went through the document. This text box allowed seven lines of text to be displayed (see Figure 3.1). Limiting the size of the text box achieved a couple of goals. First, it takes students less than one minute to read the approximately 77 words contained within the text box. This provides us with a time constraint against which we can judge possible behaviours that might occur. For example, although not directly used in the analysis, this could be used to determine if the individual was distracted from the task at hand if several minutes have gone by and no interactions with the system have occurred. Second, it provides a means to determine how much time and how quickly the student reads the portions of the document that contained the answers to the various questions. As the user scrolls down the document, they are able to select different questions that

they think might be pertinent to what they are currently reading. The questions could be selected in any order and any text the students had entered into the answer text box was saved and displayed when the corresponding question was selected. None of the participants were observed, nor reported, to have any difficulty with operating the interface.

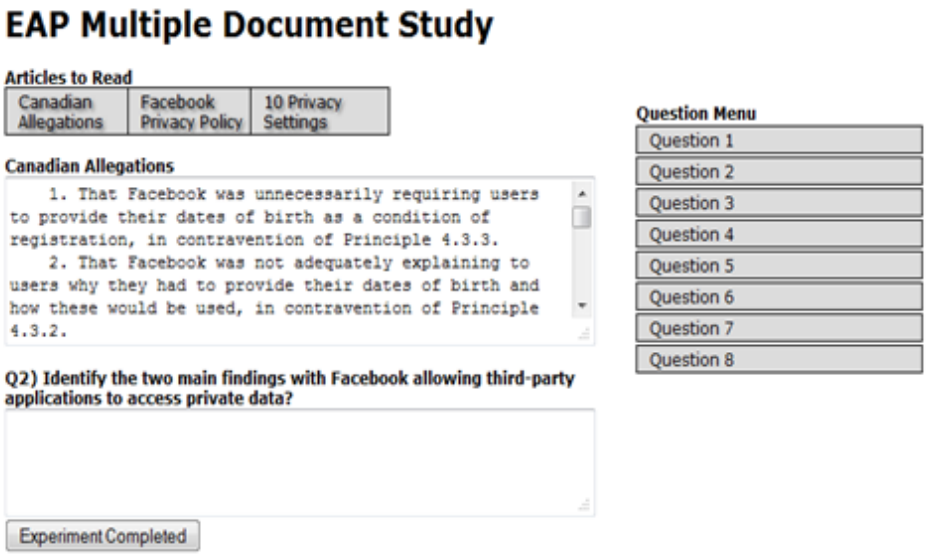


Figure 3.1 Screen Capture of Interface

The documents chosen were documents that most of the participants were unfamiliar with. For example, in Experiment 1 we used Facebook's Privacy Policy document and a corresponding document outlining Canadian Allegations about Facebook's privacy policy. Most of the Grade 12 English Adult Education students were not familiar with these documents. Anecdotally, most of the participants commented that they learned things about the privacy settings inside Facebook that they, as avid Facebook users, did not previously know about. The questions were chosen to reflect a variety of levels in Bloom's Taxonomy by using Bloom's Taxonomy Action Verbs [8] [9] (refer to Appendix 4 for more information on Bloom's Taxonomy). All the questions were present on the screen at all times and could be selected in any order by the student. They were in a specific order on the screen that was randomly chosen before the experiment began, and then remained the same for all students throughout the experiment.

3.2 Data and Granularity Issues within Timestamp Data

Timestamp data without any preprocessing will not likely yield interesting information without first being converted into a more useable form. This will involve a number of steps with our data. First, the total amount of time that occurs between the individual events is calculated providing a starting point from where we can interpret the data. The timestamp for each event is used to calculate the amount of time that a student spent prior to moving onto the next action. These actions could be mouse click, keyboard button press, or mouse scroll wheel event. The difference in time between events is used to determine the duration of each event that provides the basis for further refinement of the data. For example, a student using our interface can read a document. We can infer this from the various buttons that are clicked (scroll bar) or the mouse-wheel being moved in the reading box. Second, we track which document the student is currently reading, their current position within the document, and we also record the button press of the scroll-bar or the mouse wheel event and use the durations of these events to determine the type of reading that is taking place. If the duration between these events is extremely short, i.e. milliseconds, then we can make the assumption that they are not performing a slow methodical style of reading. Rather, we can assume that they are performing more of a scrolling type action since they are moving through the text of the document very quickly. The duration of the event when combined with the type of event allows us to create further metadata that will be useful in our analysis. For example, a mouse wheel event that has a duration of more than 5 seconds before the next mouse wheel event is recorded (with no other events occurring between these two events) allows us to assume that they are reading. Furthermore, this will allow us to further categorize the type of behaviour that is taking place, such as scanning, scrolling and reading. The time cutoffs used to distinguish reading from scanning from scrolling fit with other document navigation research [90] [91] [92]. Any time between events greater than five seconds was classified as *reading*. Any time greater than two seconds but less than five seconds was classified as *scanning* and any time less than two seconds was classified as *scrolling*. The reading time also encompassed time that the participant spent thinking about the answer before moving onto performing more reading. This time does not include any of the keystroke data that was captured while the student was answering the questions. Although this data was captured, in order to reduce the number of variables that existed in our analysis, the keystroke data was not included in the analysis for Experiments 1 and 2. The issues around how we interpret reading,

scanning and scrolling as well as the ratios and clustering will be discussed in more detail in Chapter 4.

3.3 Reading, Scanning and Scrolling Ratios

Individual time-stamped data does not provide enough information by itself to be able to draw out any specific conclusions from the data about reading comprehension. Therefore, we created an aggregate value that is composed of ratios related to reading, scanning and scrolling behaviours from the values found in the data. The ratios for the reading, scanning and scrolling were calculated by adding up the total amount of time the student spent on each question in the experiment. The total time for each of the reading, scanning, and scrolling events for each individual question was calculated and divided by the total time spent on the question. This means that each of the ratios for a specific question could be spread across multiple documents and would include actions such as re-reading the same portion of a document.

Figure 3.2 shows the ratios for question 1 on a particular experiment. Notice that the ratios do not add up to 1.0 since the time spent typing is omitted from being displayed and was not used in the analysis for Experiment 1 and 2. The reading, scanning, and scrolling ratios were then normalized, excluding the typing times.

```
Question: Q1
Reading Ratio: 0.29831353348262296
Scanning Ratio: 0.1502890904720335
Scrolling Ratio: 0.18013436064827115
```

Figure 3.2 Reading, Scanning and Scrolling Ratio

There will be times in our calculations where different levels of granularity will be used with respect to the ratios. For each question, the answers that the students are expected to provide can be found within the body of the document(s). For cognitively simple questions (lower level of Bloom's Taxonomy), the answer is often found at a specific location within a single document. For more cognitively difficult questions (higher levels of Bloom's Taxonomy), the components of the answers are often found at multiple locations within multiple documents. Our coarse-grained ratio calculation includes all of the time spent working on a particular

question irrespective of location within a particular article or the number of articles read. However, there are times in our analysis where a more fine-grained ratio calculation is required. This involves looking at the reading, scanning and scrolling ratios over the portion(s) of the document(s) where the specific answer is located. Since we know the position(s) within the document(s) that the student is reading, the location(s) within the document(s) that contained the answer, and the total duration of the student over the answer, we can look at how the student was reading, scanning and scrolling over the portion(s) of the article(s) that contained the answer. The coarse-grained ratios for the entire question can be different when compared to the fine-grained ratios found over the portion(s) of the article(s) that contain the answer.

These two levels of ratio granularity (computed over one or more complete documents vs computed over relevant portions of a single document) can lead to some important implications. When the ratios are combined with the level of difficulty for each question, as determined by Bloom's taxonomy, it is possible to tie student reading behaviour to the difficulty of the task. If a studious student scores well on some answer, their choice to scan a document that seems irrelevant versus reading the document in its entirety demonstrates a good use of scanning behaviour. If they had elected to thoroughly read every document and answer all the questions they would have consumed a larger amount of time to complete the task. This may not be a problem in some instances; however, there are always situations in life and work where time is constrained. This indicates that a student should develop or learn multiple reading styles as an important life-long learning tool. These two levels of granularity for the ratios will be more fully explored in Chapter 4.

3.4 Educational Taxonomy Granularity

One of the major contributions of this thesis is the discovery that categorizing the questions students are answering according to Bloom's Taxonomy (described in Appendix 4) is necessary to find significant patterns within the reading, scanning and scrolling clusters. We have broken the questions down into what we call "low-level" Bloom questions and "high-level" Bloom questions. These categorizations are often found in literature surrounding Bloom's Taxonomy [93] and are defined as follows:

- Low Level Bloom categories – Level 1 through 3 inclusive

- High Level Bloom categories – Level 4 through 6 inclusive

Without Bloom's Taxonomy, we were not able to find any statistically significant results. However, the accuracy of the grades is not at a fine-grained level of percentages. Rather, the clusters are only predictive of more coarse-grained letter grades such as A, B, C, D, F. However, Marzano's Taxonomy (described in Appendix 4) contains two levels of granularity with respect to the level of cognitive difficulty compared to Bloom's single level taxonomy. Like Bloom's Taxonomy, Marzano's Taxonomy contains main categories that are classified according to levels of cognitive difficulty: Bloom's Taxonomy contains 6 levels, Marzano's Taxonomy contains four levels. However, each of the four main levels in Marzano's Taxonomy also contains several subcategorizations resulting in a total of 14 different categorizations. Bloom's Taxonomy does not have these subcategories. This results in two levels of granularity for the educational taxonomies. One of the purposes of Experiment 2 is to determine if more levels of categorization will allow for more fine-grained prediction of grades. Ideally, we would like to find percentage grades. However, if we were able to determine letter grades with both + and – categories, i.e. A+ or C-, this would still be an improvement. In fact, we have an opportunity to compare Bloom's Taxonomy to Marzano's four main categories as well as comparing Bloom's Taxonomy to Marzano's subcategories to determine if Marzano's Taxonomy provides a more fine-grained prediction of grades.

3.5 Questions for Students to Answer

The questions that are used in all the experiments were categorized according to Bloom's Taxonomy Action Verbs method [8]. Questions were scored according to a rubric for each question, according to the principles laid out in [94]. This also allowed us to categorize each question to the appropriate level of Bloom's taxonomy. The rubric was revised a couple of times to take into account the various types of answers that were submitted during the beta testing phase of the system. For the lower Bloom levels, the answers generally came from one direct location within a document and so the scoring was fairly simple. For the higher level Bloom questions, information from multiple sources was expected. It was also expected that the students would bring their own prior knowledge to bear on the answer. The grading for Experiment 1 and 2 was done by myself, but completely independently of the data mining and

analysis that were done later. For experiment 3, the two teachers teaching the Adult Education English course marked the questions. The teachers only marked those students who were enrolled in their respective courses. The rubrics for each of the experiments can be found in the appendices.

Next, we will describe experiment 1 in more detail, how the study was performed and the corresponding results. Experiment 1 can be seen as a template for experiment 2 and 3. Although experiment 2 and 3 did ask different questions, the methodology, interface and experimental design follow from and build upon experiment 1.

CHAPTER 4
EXPERIMENT 1

In this chapter, we will discuss the first experiment we carried out to see if we could find, through data mining, pedagogically interesting patterns in students' reading behaviour while carrying out educational tasks. Data mining can extract information from a data set and transform it into a usable form [95]. Given we can capture a large amount of student interaction data from a student working on some educational objective within a computerized environment, it stands to reason that there will be useful information that can be extracted from this data. In particular, how a student reads the content should provide some insights into how well they learn the content. However, given the ability of data mining algorithms to find unexpected patterns, it is in our best interest to capture as much information about the student interaction data as possible within the constraints of a web-based student content learning system. With this in mind, we designed our experiments to capture the necessary reading information as described in our introduction along with all the mouse and keyboard information we could. Once the data is collected it is subjected to the data mining process and pedagogically interesting information extracted, analyzed and reported.

Section 4.1 will discuss the experiment along with the design of the interface and the data captured that is used as the basis for all the experiments. Section 4.2 will discuss the results of the first experiment spending considerable time on the different types of reading behaviour patterns and their predictability with respect to student success.

4.1 Experiment 1: The Study

Our first experiment was designed to look for patterns of student behaviour in a reading comprehension task. In fact, this experimental design was the basis of all the experiments that were conducted in this experimental program. Students interacted with a learning environment designed to emulate hypermedia courses offered in post-secondary institutions where written content is presented online along with questions about that content. The students could view the content and/or questions in any order or manner they chose with no constraints applied to their

interaction with the system. This can result in numerous types of student interactions being recorded as the students use the system.

There are a large number of possible strategies that a student could perform such as:

1. A student could move methodically through the article(s) reading them first and then looks through each question before determining the order in which they are going to answer the questions.
2. A student could read through the questions prior to reading the article(s) so that they know what to look for as they are reading.
3. A student could begin to read an article and then flip through some questions and then back to read the article and so on.
4. A student could answer all of the questions without ever reading any articles. And so on…

This type of environment provided the fewest types of constraints possible in order to try to observe as many different patterns as possible in students' reading behaviour.

Experiment 1 was broken into two parts. Experiment 1A consisted of one document and a set of questions based upon that document. Experiment 1B contained three documents and the questions based upon those documents. For experiment 1A the document chosen contained information that the majority of the participants would not have prior to the experiment. The document was a fairly technical document based upon Canadian Privacy law as it applies to Facebook. This was for research convenience to provide a novel document for the participants to read. For experiment 1B a total of three documents were used. The first document was the same document used in experiment 1A while the second and third documents were new. The second document consisted of instructions on how to implement advanced privacy features not commonly used within Facebook and the third document was a high level overview of the privacy settings used within Facebook. These documents appear in Appendix 1.

Some of the questions used in Experiment 1 (where the articles were about social media and privacy) were as follows:

- Discuss if Facebook collects personal information from sources other than Facebook.
    - This question is categorized as Bloom level 2 since there was a section in the Canadian Privacy article that dealt directly with this topic and the student was required to recall the answer.
- Critique Facebook's use of third party application
    - This question is categorized as Bloom level 5 since the student must put together an argument based on material found in the three different articles that were posted.
- Choose a side and debate if the Age Policy for Facebook usage is fair.
    - This question is categorized as Bloom level 5 since the student must put together an argument both from the information found in the three articles plus integrated with their personal views on Age Policy.

The full set of questions and their Bloom levels for Experiments 1A and 1B appear in Appendix 1. These questions were created using Bloom's Taxonomy of Action Verbs to create questions at specific levels of Bloom's Taxonomy.

For Experiment 1A, there was a single document, four questions, and all the questions were at the lower levels of Bloom's Taxonomy. The participants were given 30 minutes in which to answer the questions, including the time spent reading the document. Experiment 1B provided the students with two more documents in addition to the first document, and eight questions, at both low levels and high levels of Bloom. The higher levels of Bloom's Taxonomy require synthesis / creation and evaluation and so more information and documents were needed to allow for these requirements. Again, the answers to the high level questions could be found within the documents provided. However, in order to fully answer the higher level questions, information from more than one document was required. For experiment 1B, 90 minutes were allotted as the questions were more difficult and there were three documents that needed to be read to generate complete answers. One question was a repeated question from Experiment 1A and a second question was new but based solely on the information found in that first document. The remaining six questions were new. Participants who are new to the repeated question should behave like those who are seeing the question for the first time. For those who are seeing the

question for a second time, we should see the participant moving quickly to relocate the answer. This gives us the potential to view the first problem being solved and then subsequent problem solving. We left this analysis for future work, where we would like to explore the issues of recall and memory.

4.2 The Participants

The participants were adult students enrolled in a grade 12 Saskatchewan Institute of Applied Science and Technology (SIAST) Adult Education English course. There were 17 participants for Experiment 1A and 11 for Experiment 1B with an average age of 26. Since the amount of time required to participate in both parts might be a factor in participant involvement, both 1A and 1B were designed so that they could be run separately and using different participants depending on the participants' wishes. The participants for 1B were a subset of those involved in 1A. The students were asked to participate in 1A and only those with enough motivation stayed on to complete 1B. In the actual running of the experiment, the majority of the participants moved from experiment 1A right into experiment 1B with no delay.

The 28 participants generated over 8500 events in total from both 1A and 1B, events such as the mouse clicking on a specific button or object, the pressing of a key on a keyboard, and mouse wheel scrolling. Each event was time-stamped with the user-id, event-id, current question-id, current document-id, and position within the current document. This allowed us to determine what task/question the student was currently working on, which document they were working on, where in the document they were, and what button/keystroke they had just pressed. For example, if the student turned the scroll wheel of the mouse to move down in the document, we could then determine from the time-stamp data and the position data how quickly they moved and what material they were reading. With this information, we could begin to understand student behaviour as they work at completing the various questions.

4.2.1 The Instructions

The students were given a simple set of instructions along with a short overview of the interface. The instructions were that they were to read the documents provided an answer the questions that are based on the documents they just read. The overview of the interface involved ensuring that the students knew where the documents were located, where the questions were

located, where the text of the document would appear and where they could type in their answer and then to press the submit button when they were finished the experiment. The students were told that the system would record how they read and how they answered the questions (see Appendix 1 for full consent form). They were not given any instructions on interaction methods they should use to complete the experiment. In keeping with trace methodology approaches [4], all the interactions with the content and questions were recorded and time-stamped. These would include events such as mouse click, mouse wheel, which item was clicked or selected and so on.

4.3 The Procedure

The participants were given a consent form and experimental overview data sheet to review and sign before the experiment began. The participants were then directed to an online login form where they created their online id and password. This would allow them to login and should they decide to do so at any point to remove themselves from the experiment. If this was the first time the participants had logged in, then they were taken to the optional survey where they could enter some demographic data that could be used in the analysis of the experimental results. If the participants had previously logged in, then they were taken directly to the experiment. Once all the participants were at the main interface screen, a short demonstration of the features of the interface was provided by the experimenter so that they would be familiar and able to operate the interface. The participants were then allowed to complete all the questions and read the documents in any order they wished until the time for the individual experiment had expired. The participants were asked to save the work they had completed and were thanked as they left.

4.3.1 Reading, Scanning and Scrolling

In order to determine the kind of reading the students were doing, the timestamp data was processed so that reading, scanning and scrolling navigation times could be calculated for each interaction/event. In the 8500 events captured across the 28 participants, only 13 events had a time greater than two minutes and an additional 20 events had a time greater than one minute before another event was performed. This gives us a total of 33 events that had a time greater than one minute between events. Given the time it takes to read the content in the textbox

(described in Chapter 3), the total time between events including the reading and thinking times, was not a large enough percentage of the time spent to warrant separate classification.

The ratios for the reading, scanning and scrolling were calculated by adding up the total amount of time, excluding keyboard times, the student spent on each question in the experiment. These reading, scanning and scrolling ratios at a coarse-grained granularity, as described in Chapter 3, provided us with a behavioral picture of how a student is reading as they worked on a particular problem. As a result, they were perfect inputs to a clustering algorithm to find out if there are groups of similar behaviours that occur as a student works on a problem.

4.3.2 The Clustering Algorithm

In order to see if there were students who behaved similarly for different levels of difficulty, we implemented the Forgy method for K-Means clustering for d=3 dimensions and k=4 [12]. The three dimensions that we are looking at are the reading, scanning, and scrolling ratios over the whole document, as described above. Hammerly et al. [12] demonstrated that the Forgy method, also known as Lloyd's algorithm [96], was the preferred method for initializing the standard K-Means clustering algorithm. After some preliminary exploration of the data, the desired number of clusters k = 4 were chosen. More than 4 clusters produced some clusters where there were too few items to be statistically analyzed. Since the algorithm randomly chooses its centroid points, there is no researcher bias entering into the initial sets of clusters that were created. In order to find as many interesting clusters as we could, the Forgy K-Means algorithm was iterated multiple times. We defined interesting clusters as those clusters associated with positive or negative reading, scanning or scrolling behaviours. A positive behaviour is defined as a behaviour that results in a good grade on the question being answered. A negative behaviour is defined as a behaviour that results in a poor grade. Those clusters that presented with both positive and negative behaviours were deemed less interesting. Conversely, any cluster that tended to result in either a distinctly positive or a distinctly negative behaviour was considered interesting. Each time an interesting cluster was found, the centroid was recorded. Once multiple interesting centroids were found, the most interesting centroid found was hard coded as a starting centroid, where the most interesting cluster is the one whose composition was most uniformly positive or most uniformly negative. The hard coding removes one of the

random initializations from the Forgy initialization and inserts the most interesting centroid in its stead.

For example, the experiment started with k = 4 random clusters in the initialization. The most interesting cluster was found in the first iteration, and hard coded. The algorithm was run again with one hard coded centroid and three randomly chosen centroids to see how the other random clusters interacted since how the cluster is initialized is known to have an effect on how the other clusters form [12]. If a new cluster was discovered that was more interesting than a previous closely related centroid, the old centroid was removed in favor of the new centroid. We defined a cluster to be interesting if it was predictive of a student's grade. For example, if 88% of a cluster contained A or B letter grade. If no more centroids were discovered that were more interesting than the hard coded centroid, then the second most interesting centroid was hard coded and the remaining two centroids were left random and the above process was repeated, now with two hard coded centroids. A third hard coded cluster was added in accordance with the above procedure and the process was performed again until all four of the initialization centroids were hard coded. These 4 clusters would become the final hard coded clusters that we used in the experiment.

4.3.3 Reading Cluster Types Classified

The following clusters proved to be statistically interesting with respect to the Bloom level:

- Light Reading Cluster: 50% reading: 30% scanning: 20% scrolling (50,30,20)
- Light Medium Reading Cluster: (60,30,10)
- Heavy Medium Reading Cluster: (70,20,10)
- Heavy Reading Cluster: (80,10,10)

Two other clusters that showed up in our repetitive clustering section of the analysis, mentioned previously, were Medium Scrolling (20,20,60) and Medium Scanning (20,60,20). These clusters contained too few data points to be included in any statistical analysis that was performed. However, they did show up as a unique set of clusters consistently over the multiple iterations and so will merit some attention in future experiments.

4.4 Results

The clustering algorithm determined that the data points (ratios) belonged to unique clusters, but the question remains are these clusters statistically different from one another? Since our sample size was a total of 28, we initially combined all of our clustering results together to see if there were any significant differences between the clusters that we found. An ANOVA was performed on the clusters to see if a statistically significant relationship could be found between the different reading behaviours as clustered by k-means. The ANOVA tests were performed at the $\alpha = 0.05$ level. Table 4.1 contains a row for All Levels Combined. The numbers in this row indicate that there were no statistically significant differences found among the different clusters when the questions were not distinguished by their Bloom level.

Bloom Level	**F**	**P**	**F-Critical**
1	*79.94	3.14E-16	2.86
2	*39.31	3.74E-11	2.88
3	*147.93	4.80E-11	3.63
5	*25.56	0.000029	3.59
6	*50.77	0.000385	5.99
All Levels Combined	1.40	0.25	2.68

Table 4.1 One way ANOVA for Bloom Levels (* indicates statistical significance)

We then factored in the Bloom Level of each question into the ANOVA. Specifically, we combined all the questions with the same Bloom level together to see if the level of cognitive difficulty had an impact on the differences between clusters. Questions at Bloom levels 1,2,3,5, and 6 were available in this experiment. There were no Bloom level four questions. The null hypothesis used for these tests is that the means for each of the clusters does not vary according to the Bloom level that is being tested. In other words, the reading, scanning and scrolling means should be similar for all the clusters found by k-means. Table 4.1 shows that the differences found between the clusters for each of the Bloom levels were not due to random chance. The p-values indicate that, in all but two cases, there is a really small chance of getting these results if no real difference between the groups exists. This indicates that the clusters into which the students' reading, scanning and scrolling behaviours fall are significantly different from one

another as it relates to the level of the question according to Bloom's Taxonomy. For example, those students who were classified as Light Readers (50:30:20) based on the reading, scanning and scrolling ratios for questions categorized as Bloom level 1 were significantly different from students who were classified as Light Medium Readers (60:30:10) for questions at the same Bloom level. However, this information is only available to us after we have performed a Tukey-Kramer statistical test to locate where these differences between clusters occur. Table 4.2, below, will show this in more detail. The ANOVA itself can only accept or reject the hypothesis that the clusters are significantly different from each other. It is only the Tukey-Kramer analysis that can make an exact determination of which particular cluster is significantly different from another particular cluster; the ANOVA can only tell us that there is a significant difference between the clusters in the analysis. Further analysis, discussed later, is needed in order to see which of the clusters are significantly different from each other.

As previously mentioned, we defined clusters as interesting if they provided some predictive insights with respect to student success as defined by a grade. More specifically, we are interested in matching the student's reading, scanning and scrolling behaviour to a corresponding grade that is tied to a specific level of Bloom's Taxonomy. The more accurately a cluster was able to determine a specific grade, the more interesting it was determined to be. However, the majority of clusters often contained more than one grade. The following sections will show the mapping between clusters and grades and explain the various techniques and methods used to increase predictive accuracy of a student's grades (both good and bad) based on the cluster they belong to and the difficulty of the question they are working on.

4.4.1 Initial Predictive Clusters

Although inclusion in a cluster does not completely predict scores, it is indicative of overall performance for a given question. For example, take question 3 (see Appendix 1) in Experiment 1A (Bloom Level 1 with a single document). This question was designed to force the students to scan through the document as they needed to count the number of instances that a certain event (in this case a successful appeal on a complaint about Facebook to the Canadian Privacy Commission) occurred in the document. This, in turn, led to serious issues of completing the task in time. Working under tight time constraints is often required in academia and the workplace. 100% of the students in the Light Reading (50,30,20) cluster, which was

proportionately higher in scanning and scrolling times, achieved full marks or close to full marks. Correspondingly, those students in the Heavy Reading (80,10,10) cluster scored no better than 25% with over half of the students in the cluster scoring 0%. Since the source materials were present for the duration of the experiment and there were time constraints, the Heavy Reading strategy is not the best strategy to be used in this situation.

Cluster	**Bloom**	**Grade**	**% Accurately Predicted**
Light Reading	1	A, B, F	(Grade >= B) 60%
Medium Light Reading	1	A, B,F	(Grade >= B) 55%
Medium Heavy Reading	1	A, F	(Grade = F) 77%
Heavy Reading	1	F	100%
Light Reading	2	D, F	(Grade = F) 83%
Medium Light Reading	2	A,B,C,F	(Grade >= C) 55%
Medium Heavy Reading	2	A, F	(Grade = F) 88%
Heavy Reading	2	A, F	(Grade = F) 80%
Light Reading	3	F	(Grade = F) 100%
Medium Light Reading	3	F	(Grade = F) 100%
Medium Heavy Reading	3	A, F	(Grade = F) 75%
Heavy Reading	3	A, D, F	(Grade <= D) 86%
Light Reading	5	F	100%
Medium Light Reading	5	F	100%
Medium Heavy Reading	5	F	100%
Heavy Reading	5	A,B,F	(Grade = F) 86%
Light Reading	6	F	100%
Medium Light Reading	6	N/A	N/A
Medium Heavy Reading	6	N/A	N/A
Heavy Reading	6	A,C,F	(Grade <=C) 60%

Table 4.2 Experiment 1A and 1B Grade Prediction by Cluster Type

In order to determine if the predictive nature of a cluster inclusion (Light Reading Cluster predictive of good grades for question 3) transfers to more than one question further analysis is

required. Table 4.2 shows the mapping between letter grades and a specific level of Bloom's Taxonomy for each of the clusters. For each level of Bloom, all of the questions for that level were combined to create table above. For example, in Experiment 1A and 1B there were 4 different questions at Bloom Level 1. The % Accurately Predicted column shows the highest percentage of a specific letter grade(s) contained in that cluster for all of the questions at the specific Bloom Level. In keeping with the question 3 example described above, Bloom Level 1 Light Reading is only able to predict grades greater than or equal to a B 60% of the time. Whereas, for question 3 specifically, it was able to perform this task 100% of the time. You can also see that for the Light Reading cluster at Bloom Level 1 that some 40% of the students did not perform very well at all. This is what most teachers would expect from some assignment. As can be seen, often there is more than one grade for each of the clusters at a specific Bloom level. The positional analysis described later will provide some more insights on how to increase the level of grade predictability from 60% (Bloom Level 1 Light Reading) to a higher value.

4.4.2 Heavy Reading Strategy

As the level of difficulty for the questions increased, as measured by Bloom's Taxonomy, the Heavy Reading strategy proved to be the most successful strategy. The participants could achieve better marks compared to those that chose a Light Medium Reading strategy. For example, question 6 of Experiment 1B required the participants to synthesize various thoughts and ideas about Facebook's privacy policy garnered from multiple documents into a complete argument that did not exist in any of documents (Bloom level 6). Since there is only one question at Bloom Level 6 for this experiment, we can see from Table 4.2 that the students only existed in two different clusters, the Light Reading and Heavy Reading. Those students who fell into the Light Reading cluster all received a grade of F. For those students who fell into the Heavy Reading cluster the grades were A, C, and F. To fully answer question 6, information is required from all the documents. Additionally, they must integrate what they have read into an answer that is not directly answered in any of the documents. When we analyzed how the students interacted with the documents for the Heavy Reading cluster we found the following results:

- The students who performed Heavy Reading on only one of the documents that they were required to read did not score above 30%. (Letter Grade F)

- Those students who performed Heavy Reading on two of the required documents scored no higher than 83% (Letter Grade C)
- Those that performed Heavy Reading on all the documents scored no lower than 83% and up to 100%. (Letter Grade A)
- Those students who used the Light Medium Reading strategy scored 0%. There was one student who scored 30% that used the Light Reading strategy, but their answer contained no content from any of the documents; rather they apparently used outside information gained from their previous experiences.

Another interesting finding from Table 4.2, with respect to Heavy Reading, is that across the four different questions at Bloom Level 1 all the students that fell into this cluster received a grade of F. Furthermore, if we look at the Medium Heavy Reading and the Heavy Reading clusters from the Bloom Levels 1 and 2 we see that the majority of those students obtained a grade of F.

4.4.3 Light Reading and Light Medium Reading Strategies

Recall Table 4.2, when we look at Bloom Level 3, 5 and 6 we see that the use of the Light Reading and Light Medium Reading strategies result in poor grades (F). As the level of cognitive difficulty increased, the students needed to be able to put information together from multiple sources in order to be able to obtain good grades. The strategy of scanning and scrolling through the documents no longer provided a method to obtain good grades. The participants needed to be able to recall information from a variety of sources in order to be able to fully answer the questions. Instead of using source material, possibly because they could not recall where it was or if it was present, they used incorrect information from some other source outside of the experiment. When the answers were checked manually, it was found that they had used some incorrect prior knowledge from some other source than the sources provided. It should be noted that they did not access supplementary material from either books or the Internet during this experiment, which leaves us to conclude that this information was recalled from memory.

4.4.4 How Position within the Document Affects Prediction

Recall Table 4.2 contained multiple instances where the clusters for a particular Bloom level contained more than one grade. The results in Table 4.2 are an aggregate of all the

questions for a specific Bloom level. In order to better understand how to better predict grades for a specific cluster when more than one grade is present, we further analyzed specific questions where multiple grades occurred in a given cluster. For example, question 2 (Experiment 1A, Bloom Level 1) we have the A, B, and F grades represented within the Light Reading cluster. The reading, scanning and scrolling ratios are calculated for the total time taken to answer a specific question across all of the documents. However, a person may exhibit different reading, scanning, and scrolling behaviours for different portions of a document while trying to answer that question. Since we captured the current reading position within a document with each event, we can determine the amount of time spent reading, scanning and scrolling over the portions of the document that contain the answer. For example, one student classified as Light Reading when you looked at the entire time the student spent working on question 2. However, when we looked at the total time that the student spent reading over the portion of the document that contained the answer they clustered as a Medium Heavy Reading. This demonstrates that a student's reading, scanning and scrolling ratios can differ for different portions of the document.

Cluster	**Bloom**	**Grade**	**% Accurately Predicted**
Light Reading	1	B	100%
Medium Heavy Reading	1	A	100%
Scanning / Scrolling Over Answer	1	F	100%

Table 4.3 Analysis of Reading, Scanning, and Scrolling Over the Answer for Question 2

Table 4.3 shows how the multiple grades for the Bloom Level 1 Light Reading cluster is broken down when the ratios over the portions of the document that contain the answer are recorded. This seems to suggest that the reading scanning scrolling ratios become more accurate in terms of predicting grades when you analyze the actual ratio over the portions of the document(s) that are required to fully answer the question. This implies that there are levels of granularity with respect to the ratios as they appear in the sections of the document that contain the answer compared to the ratios for the entire question.

4.4.5 Marzano's Taxonomy

From the bottom of Table 4.1, we know that without taking into account the level of the question students are answering in Bloom's Taxonomy, we don't find any significant differences

between the clusters as to their ability to predict grades. Only after we break down the clusters by the different levels of Bloom's Taxonomy do we find significant differences between the different clusters in this ability. Table 4.1 showed that there were significant differences for all of the levels of Bloom's Taxonomy. Table 4.2 shows the different grades that correspond to each of the clusters at different levels of cognitive difficulty. Table 4.3 demonstrates that in cases where there are multiple grades within a single cluster, that closer examination of the reading, scanning and scrolling ratios as they relate to where the answer is located do help provide better accuracy of predicting a grade. This poses an interesting question: Are there other similar taxonomies, such as Marzano's Taxonomy (See Appendix 4 for more details), that might provide a more fine-grained distinction between the cognitive differences for problems?

4.4.6 Tukey-Kramer Analysis: Which Clusters are Significantly Different from Each Other?

The ANOVA calculation showed in Table 4.1 that there were significant differences between the clusters but an ANOVA cannot show where those differences occur. In order to find out which clusters were significantly different from each other, a Tukey-Kramer analysis is required. A Tukey-Kramer analysis allows for post-hoc multiple pairwise comparisons of each of the clusters for groups that do not have the same number of students. The Tukey test can be thought of as a T-Test with error handling built in for multiple comparisons and the Kramer test controls for unequal means within the pair being tested. Since from the ANOVA calculation we have already accepted the hypothesis that there are differences between the clusters in predicting grades for questions at various levels of Bloom, we can now move on to a more specific analysis of the differences of the means between the clusters by comparing them to each other at each Bloom level using Tukey-Kramer. The minimum significant difference value was used to calculate the pairwise comparison to determine if the elements of the pair are significantly different from each other and to correct for multiple comparisons. The numbers in the top right hand portion of the Tables 4.4 through 4.6 show the Tukey-Kramer minimum significant differences (MSD). The numbers in the lower left corner of Tables 4.4 through 4.6 show the observed absolute value of the difference in means between each pair of groups. Those numbers in the lower left of the tables marked by an asterisk are deemed significant if they are larger than their corresponding MSD located in the top right of the table. All the values are calculated from all the questions combined for each level of Bloom.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.08311	0.07745	0.08089
60,30,10	0.16204*	-	0.07976	0.08311
70,20,10	0.2963*	0.13426*	-	0.07745
80,10,10	0.4447*	0.28268*	0.14842*	

Table 4.4 Tukey-Kramer Analysis Bloom Level 1 (* denotes significant differences)

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.21238	0.19337	0.17629
60,30,10	0.21341*	-	0.2324	0.21839
70,20,10	0.4906*	0.2772*	-	0.19995
80,10,10	0.6724*	0.459*	0.18183	-

Table 4.5 Tukey-Kramer Analysis Bloom Level 2

Table 4.4 shows that all of the clusters at Bloom level 1 were significantly different from each other. Table 4.5 shows that for Bloom level 2, there are significant differences between most of the groups except for the Medium Heavy Reading cluster and the Heavy Reading. Although the k-means algorithm clustered these reading, scrolling and scanning ratios into two different clusters, the actual Euclidean distance between the ratios in the two clusters was close. Figure 4.1 provides an example of what that might look like.

Figure 4.1 Euclidean Distance Between Two Close Clusters

Table 4.2 shows that the grades for these two clusters were both A, F with the majority of both clusters not performing well (Medium Heavy Reading, Grade = F 80%; Heavy Reading, Grade = F 88%). The similarities between grades and the Euclidean distance between the data points

make these two clusters non-significant. It was situations like this one where the clusters were close together that made us wonder if a breakdown of individual Bloom levels was the best predictor. If we could look at a more fine-grained level of cognitive difficulty, such as Marzano's Taxonomy, would we get the types of breakdowns that we found in Table 4.3?

Table 4.6 shows a similar pattern to Table 4.5 with respect to Euclidean distance. However, if we look at Table 4.2, we see that for the Medium Light Reading and the Medium Heavy Reading clusters are completely absent for Bloom Level 6. At Bloom Level 5 we see that there is no longer a significant difference between these the Medium Light Reading and the Medium Heavy Reading clusters. This may be pointing out that these clusters should be disappearing as the level of cognitive difficulty increases.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.40104	0.44838	0.39138
60,30,10	0.6789*	-	0.3063	0.21437
70,20,10	0.47*	0.20891	-	0.29353
80,10,10	0.9859*	0.30701*	0.5159*	-

Table 4.6 Tukey-Kramer Analysis Bloom Level 5

One of the major problems in this experiment was that we did not have a large enough sample size for the higher levels of Bloom as tested in part 1B. These results need to be replicated in another experiment as well as provide a larger value of N for many of these results.

4.4.7 Gabriel Comparison Interval

Since we are performing a post-hoc analysis of the data, the Gabriel Comparison Interval (GCI) provides a much more accurate measure of comparing means across a group of unplanned comparisons that take place in a Tukey-Kramer analysis. The ANOVA and Tukey-Kramer analysis show where there are statistical differences between the different clusters but they do not show how accurate the differences are. This can be a huge problem when there is analysis that can have varying sample sizes. The GCI is a statistical measure similar to the standard error of means or 95% confidence intervals. Standard deviations do not take unequal group sizes into consideration and tend to be a poor report of error range. GCI takes into account the number of observations within the group, the number of groups, and the desired probability level (alpha) in

its calculation. Since the probability of a Type 1 error increases with the number of tests, the GCI takes this into account in its calculation. As the number of tests increase the GCI becomes more conservative. This means that groups with smaller values of N will have intervals that tend to be larger than those groups with a larger N. Any of the intervals that overlap between groups are considered to not be statistically significant.

Figure 4.2 demonstrates that all of the groups are significantly different from each other since no groups overlap. The X-axis of the figure are the individual clusters and the Y-axis is the means with upper and lower confidence interval bars. Since the confidence intervals for each of the groups is relatively small, we can see that our N for each group is a statistically good size. The Y-axis values in the middle are the means for each cluster and the upper and lower confidence interval show the variation within the cluster accounted for sample size, number of groups and the alpha level from the ANOVA.

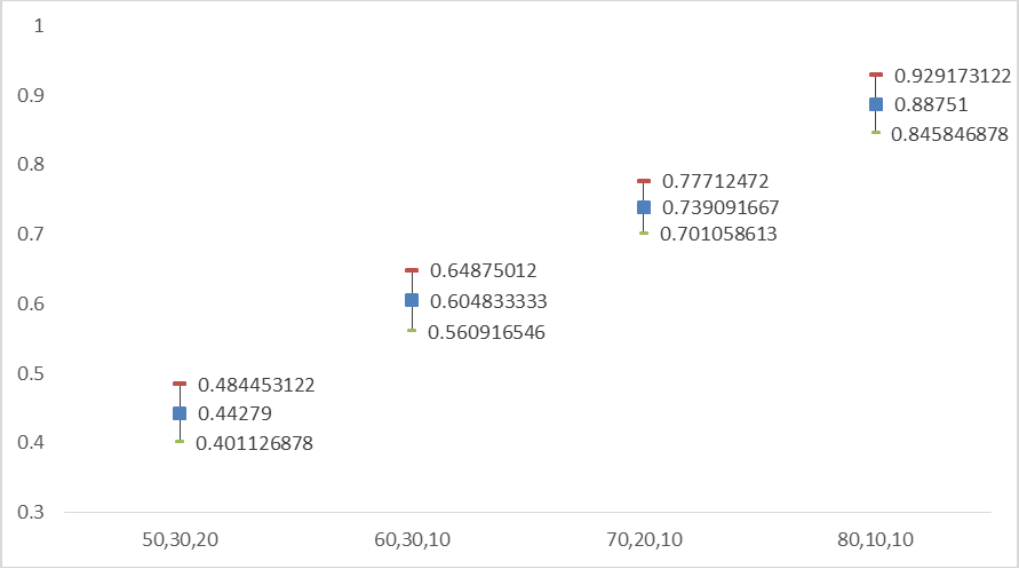


Figure 4.2 Gabriel Comparison Interval for Bloom Level 1

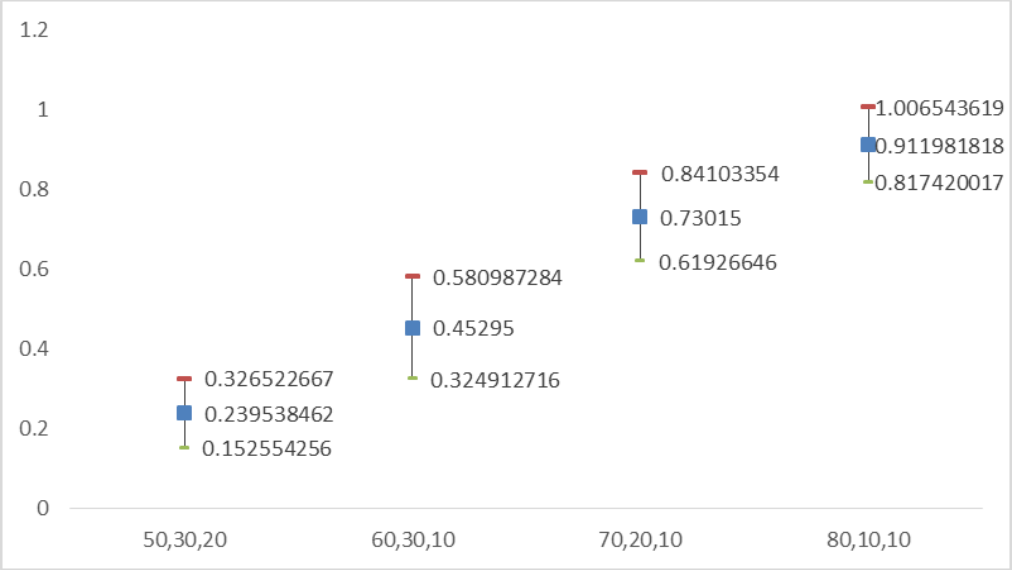


Figure 4.3 Gabriel Comparison Interval for Bloom Level 2

Figure 4.3 shows that the 60,30,10 (Medium Light Reading) cluster has a smaller N value than does the 70,20,10 (Medium Heavy Reading) cluster. Since the 70,20,10 cluster overlaps with the 80,10,10 (Heavy Reading) cluster there is no significant difference between these two clusters at Bloom level 2.

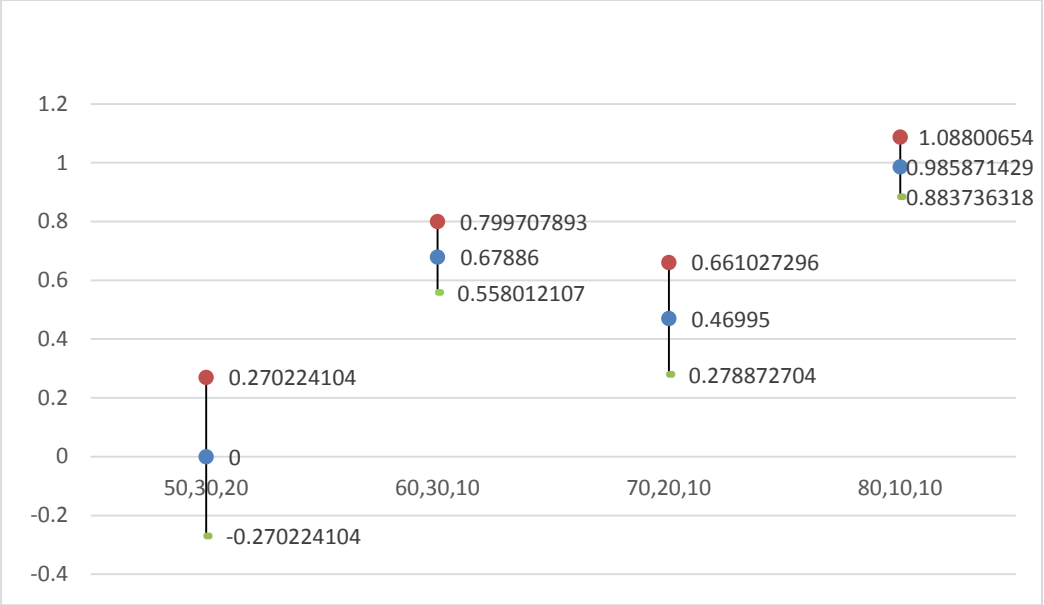


Figure 4.4 Gabriel Comparison Interval for Bloom Level 5

Figure 4.4 shows that the 70,20,10 cluster does not have a large N and that its confidence interval overlaps with the 60,30,10 cluster, making those two clusters not significantly different from each other. Furthermore, it points out that the Light Reading (50,30,20) cluster only had grades of zero. The large confidence intervals point out that not many people were in this cluster. The GCI helps pull together the Tukey-Kramer analysis by showing where our N was small due to the students not answering the questions. It also helps point out where a student received a zero versus not answering the question at all. In the Tukey-Kramer analysis for Bloom Level 5, there are significant differences between the Light Reading (50,30,20) and the Medium Light Reading, and the Heavy Reading clusters. But the GCI shows that the person received a zero, so we could ignore the Light Reading cluster without losing too much information.

4.4.8 How Level of Cognitive Difficulty Interacts with Reading Types

Next we analyzed how the level of cognitive difficulty (the Bloom level) interacts with the various reading clusters. Figure 4.6 shows how the different reading styles (clusters) were used across the various levels of Bloom's Taxonomy. For example, the Light Reading (50,30,20) behaviour was not found in any questions above Bloom level 3. This seems to indicate that Light Reading behaviour is not conducive to the more cognitively difficult tasks. This does not mean that Light Reading does not occur at the higher Bloom levels however; the ratios are an aggregate of all reading behaviour for a particular Bloom level. It is this aggregated ratio that is indicative of behaviour. The Heavy Medium Reading cluster had only 2 instances in questions above Bloom level 3. The decreasing use of Heavy Medium Reading as the Bloom level of difficulty increases shows that some of the students adopted a heavier reading behaviour compared to their use of the Heavy Medium Reading behaviour at the lower Bloom levels. They gave up the Heavy Medium Reading strategy for the Heavy Reading strategy used more in the higher Bloom levels. The Heavy Reading cluster was found at each of the Bloom levels. As the Bloom levels increase in difficulty, the amount of Heavy Reading increases until all students (except one) are Heavy reading at Bloom level 6. Correspondingly, the Light Reading cluster that contains more scrolling and scanning decreased as the Bloom level increased. This seems to confirm our earlier findings that different strategies are appropriate for different Bloom levels.

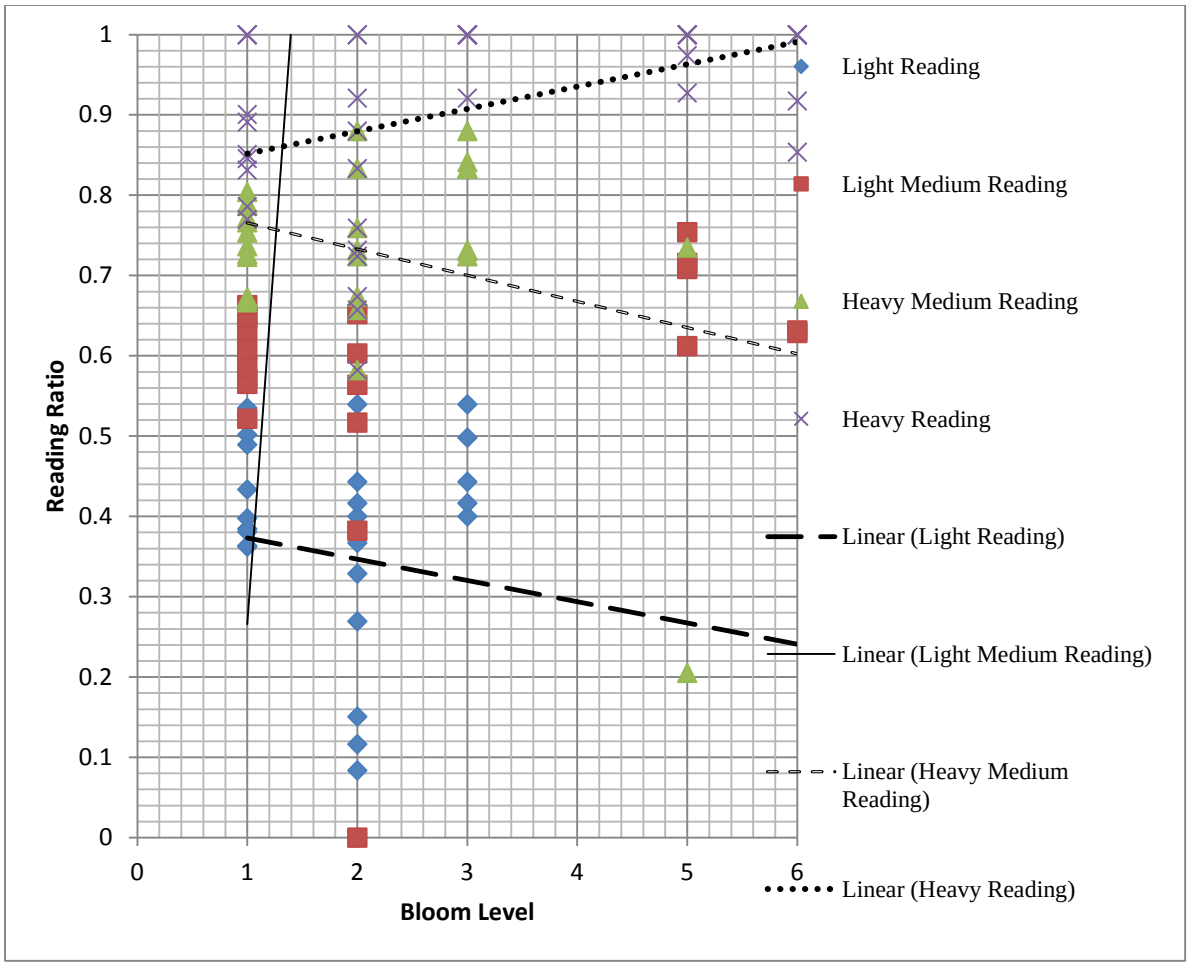


Figure 4.5 Graph of Reading Ratio vs Bloom Level

At Bloom level 6 only two strategies are used: the (60,30,10) Light Medium Reading and the (80,10,10) Heavy Reading strategies. Although the students' inclusion in the Heavy Reading cluster was a good indicator of higher scores, there was still a lot of variance in the grades found within the Heavy Reading cluster for Bloom level 6. The best predictor of scores within the cluster was the ratio of reading time spent when a student was focused on the position in the various documents that contained the material necessary for the answers. This helped to identify those students who merely used their own unsupported opinions to answer questions versus those students that used information from the articles to support their answer.

4.5 Summary of Results

There are two main results from our first experiment. The first is that Bloom's Taxonomy should be incorporated into the analysis of reading behaviour to draw out pedagogical meaning. Clustering algorithms will cluster around the data even if the clusters make no pedagogical sense. However, the inclusion of Bloom's Taxonomy as an integral part of the process has allowed us to make predictions about student success as it relates to grades. Second, the further analysis found in Tables 4.2 through 4.4 demonstrated that different patterns of reading, scanning and scrolling resulted in the ability to predict good or poor grades, at a coarse-grained letter grade level, for different levels of cognitive difficulty as defined by Bloom's Taxonomy.

The statistically significant results found when Bloom's Taxonomy was incorporated into the clustering results was encouraging. However, we have to make sure that over-fitting did not occur. Further, our N of just 28 is not as large as we would like. If we keep the changes to the interface and the experimental design minimal in our next experiment, we should be able to combine the results both experiments to help increase the size of N.

Not all of the questions were completed by the participants. The value of N decreased as the level of cognitive difficulty increased. This was especially noted in those participants that kept the same heavy reading style for all of the questions they answered. As a result of this reading style, they ran out of time to complete all of the answers. Conversely, those students who changed reading styles based upon the level of cognitive difficulty of the question could complete all the questions in the time allotted. The next experiment will need to increase the size of N at the higher levels of Bloom's Taxonomy.

CHAPTER 5
EXPERIMENT 2

Experiment 2 extends our earlier work from experiment 1 [97]. In experiment 1, we ran the clustering algorithm through many iterations until we found useful clusters. This ran the risk of over-fitting our results. If our hard-coded clusters are equally predictive of student grades using different subject matter and a different set of students, the possibility of over-fitting is reduced. We also wanted to gain more insight into reading strategies when there is more than one document. Finally, we wanted to see if a more refined perspective on cognitive difficulty could provide more insights into student reading, scanning and scrolling behaviours. Marzano's Taxonomy (for a full description see Appendix 4) offers four main levels of cognitive difficulty compared to Bloom's 6 levels of cognitive difficulty. However, Marzano also offers further refinements in each of its four main difficulty levels resulting in 14 different cognitive levels of difficulty compared to Bloom's 6 levels of cognitive difficulty. Do Marzano's extra levels for categorizing questions make any real difference compared to using Bloom's Taxonomy?

5.1 The Study

Building on our first experiment, a second experiment was performed using the same environment as in experiment 1B, but with different participants and with readings on different topics. By expanding the topics and participant types we were adding to the robustness of our approach. Experiment 2 was divided into two different parts with both parts containing multiple documents and multiple questions. The first part, experiment 2A, asked four questions, all at the lower Bloom levels (defined in Chapter 3) across five documents. The participants were given 90 minutes since more reading was required. The second part, experiment 2B, asked four questions, all at the higher Bloom levels (defined in Chapter 3) across four documents. The questions for both parts were designed using the action verb methodology of Bloom's Taxonomy described in Appendix 4 and the results were graded by myself (independently of the analysis of student behaviour) with a rubric to reduce subjectivity in the grading. The consent form, documents, questions, categorizations, and rubrics for experiment 2A and 2B are all available in Appendix 2.

5.1.1. Clustering for Experiment 2

The k-means clustering algorithm used to locate the centroids in experiment 1 chooses random starting points from the data to create the clusters. This results in different clusters showing up depending on the starting point selected. For experiment 1 we performed multiple iterations of the k-means algorithm looking for those clusters that showed up most often. We then analyzed these clusters to see if there were any pedagogically interesting features within those clusters. There were four clusters whose centroids we hard-coded as the starting points that we reported on in experiment 1.

For experiment 2 the 4 hard coded clusters that were used in experiment 1 were chosen as the starting centroids for the clusters in experiment 2. We did not perform k-means clustering to find our starting centroids. To further help ensure that we were not over-fitting our results, we chose a different topic (hacking) along with a different type of participant cohort.

5.2 The Participants

There was a total of 22 participants in experiment 2A and 2B. The participants were students from a local college computer program in Saskatoon and computer graduate students from a local university. There were 15 participants for experiment 2A and 7 participants for experiment 2B. All of the participants in 2B also participated in 2A. The age of the participants ranged from 18 through to 44.

5.2.1 Instructions

The students were given a simple set of instructions along with a short overview of the interface. The instructions were that they were to read the documents provided an answer the questions that are based on the documents they just read. The overview of the interface involved ensuring that the students knew where the documents were located, where the questions were located, where the text of the document would appear and where they could type in their answer and then to press the submit button when they were finished the experiment. The students were told that the system would record how they read and how they answered the questions (see Appendix 2 for full consent form). They were not given any instructions on interaction methods they should use to complete the experiment. Participants were told they had 90 minutes to

complete the task and were allowed to start the experiment. I was present to answer questions about the interface or provide clarifications about what the questions were asking.

5.3 The Procedure

As in experiment 1B, we captured each event that the student carried out. In keeping with trace methodology approaches [4], all of the interactions with the content and questions were recorded and time-stamped. These include events such as mouse click, mouse wheel, which item was clicked or selected and so on. The time cutoffs used to differentiate between the reading, scanning and scrolling categories were the same as those used in experiment 1 [29].

The second experiment made use of the four centroids discovered from the first experiment [29] to see if the same clusters emerge as those we found in experiment 1. As in experiment 1 the null hypothesis is that the reading, scanning and scrolling means should not be significantly different between the clusters found by k-means.

Experiment 2 involved a total of 22 participants creating 55,238 events over the two parts of the experiment. The larger number of events captured by the experiment (when compared to experiment 1) has to do with the fact that the second experiment involved multiple documents for both 2A and 2B. Experiment 2A with 15 participants recorded 30,523 events and experiment 2B with seven participants recorded 21,715 events.

5.4 Results

There are two main results that we are looking for in this experiment. First, we want to confirm that the results from experiment 1 hold and that we have not over-fitted the algorithm. Furthermore, we did not test Bloom Level 4 in the first experiment. To that end, we have multiple questions at Bloom level 4 and 5 with no Bloom level 6 questions in experiment 2B. We left Bloom level 6 for the third experiment. Second, we are interested in finding out if a more fine-grained level of cognitive difficulty improves the predictability of grades compared to Bloom's Taxonomy used in experiment 1.

5.4.1 Confirmation of Experiment 1 Results

Recall that the results of experiment 1 (shown in Table 4.1) show that the students' reading, scanning and scrolling behaviours captured by the system and then clustered are

significantly different from one another when the level of the question in Bloom's Taxonomy is taken into account.

Table 5.1 shows that for experiment 2 (part 2A and 2B reported together) all the levels tested to have significant differences. This shows that the same four centrums taken from the first experiment also clustered the data from the second experiment into statistically significantly different groups.

Bloom Level	**F**	**P**	**F-Critical**
1	*23.137	1.04E-6	3.09
2	*33.245	2.47E-7	3.19
3	*21.237	.005796	6.60
4	*50.535	.000854	6.60
5	*25.128	1.18E-6	3.15

Table 5.1 One way ANOVA for each Bloom Level Experiment 2

Again as in experiment 1, inclusion in a particular cluster does not give an exact percentage grade; rather it provides a prediction of a more coarse-grained letter grade. Question 2 in experiment 2A asked for the students to recollect two pieces of information. Students in the Heavy Reading cluster when answering question 2 almost always received a failing grade, while those students who performed more scanning obtained a grade greater than 75%. Those students who performed more scanning and who did not receive passing grades did so because they misinterpreted the question, in the sense that the answer provided by the students was in no way related to the question asked. Furthermore, the answers were not contained within the text of the documents they had viewed. This probably accounts for the high scanning as they were looking for an answer that they had in mind but could not find in the text of the documents they viewed.

5.4.1.1 Grade Prediction by Cluster Type

As it was done in Experiment 1, we mapped the grades to each of the clusters for their corresponding level in Bloom's Taxonomy. Again, multiple questions could map to the same

level of Bloom's Taxonomy from both Experiment 2A and 2B. Table 5.2 shows that the results are similar to Experiment 1 (Table 4.2). For example, the Light Reading for Bloom Level 5 has multiple possible values and does not have a good % accuracy for predicting grades greater than or equal to a B. There was one question from Experiment 2A and one from Experiment 2B for Bloom Level 5. Although the overall topic for these questions was similar, the precise content that was covered was very different. For this experiment the overall subject matter was "Hacking" with the majority of participants in this study being computer science students. This

Cluster	**Bloom**	**Grade**	**% Accurately Predicted**
Light Reading	1	A, B,D, F	(Grade >= B) 77%
Medium Light Reading	1	A	100%
Medium Heavy Reading	1	A, D	(Grade = A) 80%
Heavy Reading	1	A,D,F	(Grade = A) 85%
Light Reading	2	A	100%
Medium Light Reading	2	N/A	N/A
Medium Heavy Reading	2	A	100%
Heavy Reading	2	A, D	(Grade = A) 81%
Light Reading	3	A,B	(Grade >=B) 100%
Medium Light Reading	3	N/A	N/A
Medium Heavy Reading	3	A	100%
Heavy Reading	3	A	100%
Light Reading	4	F	100%
Medium Light Reading	4	N/A	N/A
Medium Heavy Reading	4	N/A	N/A
Heavy Reading	4	A, F	(Grade = F) 75%
Light Reading	5	A,B,C,F	(Grade >= B) 55%
Medium Light Reading	5	N/A	N/A
Medium Heavy Reading	5	A,B,D	(Grade >= B) 75%
Heavy Reading	5	A,B,D	(Grade >= B) 89%

Table 5.2 Experiment 2A and 2B Grade Prediction by Cluster Type

is most likely the main reason for the grades being higher in this experiment compared to Experiment 1.

Table 5.3 shows a further analysis of the Light Reading group from Bloom Level 5 on Table 5.2. We again performed a breakdown of how the student was performing over the areas where the answers were located. Since this was a higher level Bloom question the areas over where the necessary material was located where much larger and across multiple documents. As can be seen by the results of Table 5.3, we are able to increase the accuracy of our predictions when we examine the type of reading that was taking place directly over the areas where the answer was located.

Cluster	**Bloom**	**Grade**	**% Accurately Predicted**
Light Reading	5	C	100%
Heavy Reading	5	A,B	(Grade >= B)100%
Scanning / Scrolling Over Answer	5	F	100%

Table 5.3 Positional Analysis of Bloom Level 5 Light Reading for all questions

5.4.1.2 Where the Significant Differences Occur in Experiment 2

Similar to experiment 1, a Tukey-Kramer analysis was required to determine where the significant differences occurred within the particular Bloom Level. The similarity of the results despite the differences between the participant groups and differences in the subject of the documents adds weight to the assertion that the results we obtained in the first experiment were not due to over-fitting the data but to reliability of the results. The results for the other Tukey-Kramer analysis can be found in Appendix 2. Due to the similarity of the results between experiment 1 and experiment 2, I chose to report only on those Bloom levels I did not report on in experiment 1.

The numbers in the top right hand portion of Tables 5.4 and 5.5 show the Tukey-Kramer minimum significant differences (MSD). The numbers in the lower left corner of Tables 5.4 and 5.5 show the observed absolute value of the difference in means between each pair of groups.

Those numbers in the lower left of the tables marked by an asterisk are deemed statistically significant if they are larger than their corresponding MSD located in the top right of the table.

Table 5.4 and 5.5 show the differences between the clusters for experiment 2 (2A and 2B combined). Again, we see that there are significant differences but those differences tend to be between the 50,30,20 (Light Reading) and the 80,10,10 (Heavy Reading) clusters. In the second experiment, the participants consisted primarily of individuals who are experienced and frequent computer users. This contrasts with the participants in experiment 1 who were primarily novice computer users. The participants in the second experiment tended to either perform heavy reading or the other extreme consisting of the highest scanning and scrolling ratios (Light Reading). The middle two clusters were under-represented in the second experiment. These results seem to indicate that individuals who use computers often seem to opt for either a Light Reading or a Heavy Reading strategy depending on their preference and need.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.19626	0.15202	0.13407
60,30,10	0.25348*	-	0.1896	0.17554
70,20,10	0.3588*	0.10529	-	0.12412
80,10,10	0.4651*	0.21159*	0.1063	-

Table 5.4 Tukey-Kramer Analysis Bloom Level 2 Experiment 2

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.49212	0.49212	0.21581
60,30,10	0.3534	-	0.66866	0.5015
70,20,10	0.5084*	0.155	-	0.5015
80,10,10	0.6527*	0.29936	0.14436	-

Table 5.5 Tukey-Kramer Analysis Bloom Level 5 Experiment 2

5.4.1.3 Gabriel Comparison Intervals for Bloom Levels 2 and 5

Figure 5.1 and 5.2 show the GCIs for the Tukey-Kramer results reported above. Figure 5.1 has a larger N and as a result we can see that the confidence intervals are much tighter compared to Figure 5.2. In Figure 5.1, we also notice that the confidence intervals for the 60:30:10 cluster

are larger than the other three. This indicates that fewer users used the 60:30:10 strategy compared to the other strategies. As our N increases we expect to see that there are statistically significant differences between all the clusters.

Figure 5.2 shows that there were really only two strategies predominantly used at Bloom level 5. As in the Bloom Level 6 analysis from experiment 1, those participants who performed a Heavy Reading strategy received good grades (letter grade B or greater). Those participants who chose the Light Reading strategy did not receive a passing grade. With an N of 7, one should be cautious about the interpretation of these results. However, since these results seem to corroborate the results we found in experiment 1, they have more weight than if they appeared without other substantiating evidence.

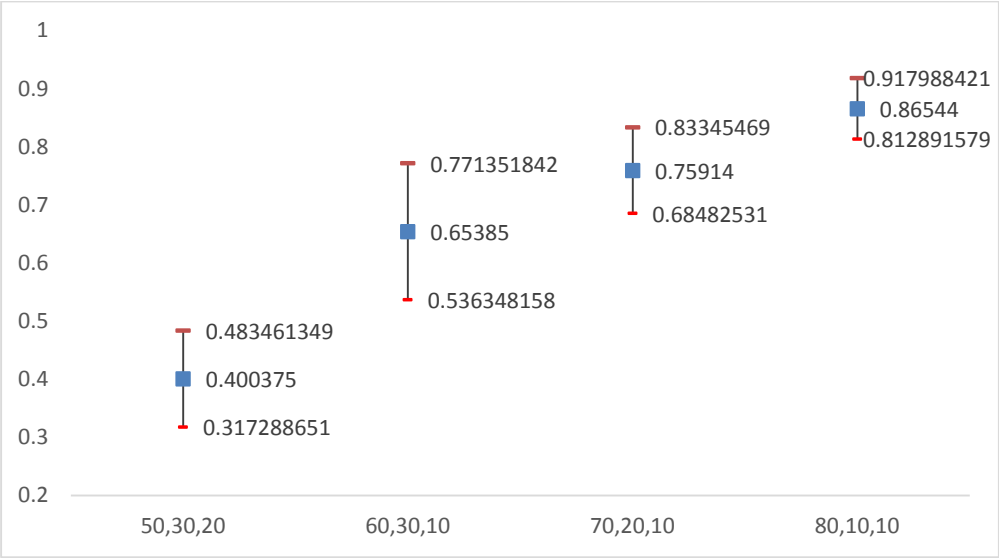


Figure 5.1 Gabriel Comparison Interval for Bloom Level 2 on Experiment 2

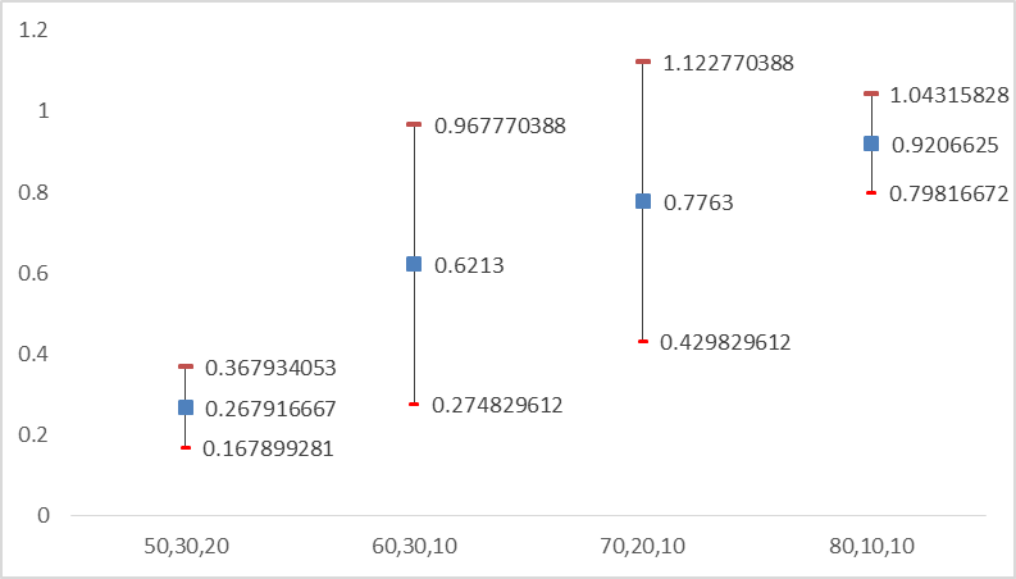


Figure 5.2 Gabriel Comparison Interval for Bloom Level 5 on Experiment 2

5.4.2 Coarse-Grain Levels of Cognitive Difficulty with Marzano's Taxonomy

As mentioned previously, we were wondering if more levels of cognitive difficulty would provide a finer grained prediction of grades. A separate analysis was performed on the data collected in experiment 2A and 2B by recasting the categorizations from Bloom's Taxonomy to Marzano's Taxonomy. This was done by mapping the questions from Bloom's Taxonomy over to the four main categories found in Marzano's Taxonomy. Another separate analysis was performed a second time, but this time we mapped from Bloom's Taxonomy to Marzano's 14 subcategories rather than the main categories in Marzano. Each of the questions was analyzed to determine its category and sub-category within Marzano's taxonomy.

We first do an analysis using Marzano at a coarse grain size to compare it to Bloom, and then move on to an analysis using the finer grained Marzano levels below.

Marzano Level	F	F-Critical
1	*120.98	2.73
2	*62.31	3.07
3	*52.71	3.91
4	0.60	3.58
All Levels Combined	1.40	2.67

Table 5.6 ANOVA for Marzano Experiment 2

Table 5.6 shows the clusters generated for each of the first 3 levels of Marzano were significantly different from the other clusters in each level for the second experiment. Level 4 of Marzano's Taxonomy did not show up as statistically significantly different from any of the other clusters. There were not sufficient numbers of participants in the experiment to obtain statistically significant values for all levels of Marzano's Taxonomy. We will not provide a Tukey-Kramer result for this table as we will be recasting both experiment 1 and experiment 2 questions together in the next section to increase the size of our N.

5.4.2.1 Recasting Both Experiment 1 and Experiment 2 to Marzano's Taxonomy

Since we retained the data from the first experiment, if we recast both the first and second experiment in terms of Marzano's Taxonomy, we increase our sample size for a given level of difficulty. When we combined the results of both experiments 1 and 2 using Marzano's Taxonomy, the clusters at Marzano Level 4 become significant F = 43.86, F-Critical = 3.00, p = 6.77E-10. The corresponding ANOVA can be found in Appendix 2.

5.4.2.2 Tukey-Kramer Analysis for Course-Grain Marzano Levels for Experiments 1 and 2

We performed a Tukey-Kramer analysis to find out where the differences were between the clusters. Table 5.7 through 5.9 show the significant differences that exist for the 4 coarse-grained Marzano levels. What is interesting is that for all the levels, there seems to be no significant difference between the 60,30,10 and the 70,20,10 clusters. We are not sure if the

more coarse-grained values of Marzano's 4 levels compared to Bloom's 5 and 6 levels are blurring the difference between the two clusters that show significant differences in Bloom or if it is the differences in the cross over between low and high reading clusters that are causing this effect.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.11024	0.09268	0.08039
60,30,10	0.27235*	-	0.11431	0.1046
70,20,10	0.3523*	0.07993	-	0.08589
80,10,10	0.5613*	0.28894*	0.20902*	-

Table 5.7 Tukey-Kramer Analysis Marzano Level 1 for Experiment 1 and 2

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.17919	0.10896	0.07835
60,30,10	0.23278*	-	0.19086	0.1752
70,20,10	0.26605*	0.03327	-	0.10225
80,10,10	0.4193*	0.18652*	0.15325*	-

Table 5.8 Tukey-Kramer Analysis Marzano Level 2 for Experiment 1 and 2

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.20153	0.12675	0.09546
60,30,10	0.3569*	-	0.2102	0.19295
70,20,10	0.4292*	0.07231	-	0.11261
80,10,10	0.6894*	0.3325*	0.2602*	-

Table 5.9 Tukey-Kramer Analysis Marzano Level 3 for Experiment 1 and 2

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.22324	0.22324	0.18434
60,30,10	0.4831*	-	0.23317	0.19624
70,20,10	0.30819*	0.17492	-	0.19624
80,10,10	0.7433*	0.26019*	0.4351*	-

Table 5.10 Tukey-Kramer Analysis Marzano Level 4 for Experiment 1 and 2

When we look at Table 5.7, the Marzano level 1 clustering we find that it offers the same patterns of predictability as we found in Bloom level 1. For the coarse-grained level of cognitive difficulty, there was no improvement in the level of predictability and it is still at the letter grade level of granularity for grades. This isn't surprising since Marzano's Taxonomy has fewer levels than Bloom's Taxonomy.

5.4.2.3 Gabriel Comparison Interval for Marzano Coarse-Grained Levels

Figure 5.3 shows the Gabriel Comparison Intervals for the Marzano analysis using combined data from experiments 1 and 2. The confidence intervals are much tighter, reflecting the higher N that resulted from combining the data.

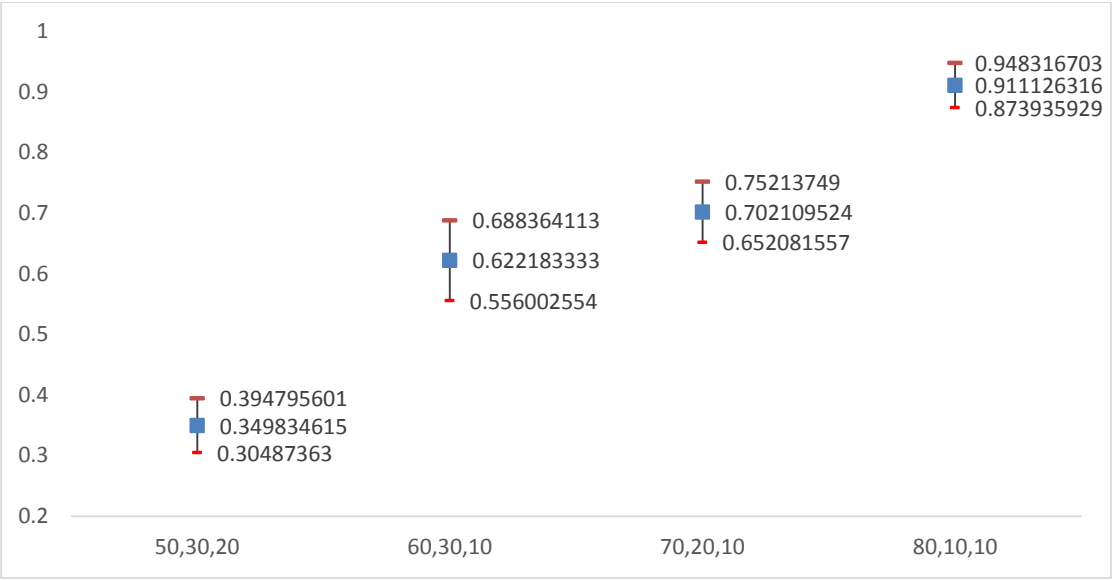


Figure 5.3 Gabriel Comparison Interval for Marzano Level 1 for Experiment 1 and 2

The overlap between the 60,30,10 and the 70,20,10 clusters shows that there are no significant differences between these clusters. The means for these two clusters are fairly close indicating that three clusters are all that we can distinguish in Marzano's Taxonomy at a coarse grain size, compared to Bloom's Taxonomy. The other GCI comparisons are similar to Figure 5.3 and can be seen in Appendix 2.

5.4.2.4 Fine-Grained Levels of Cognitive Difficulty for Marzano's Taxonomy

Marzano's cognitive domain contains four main levels that can be subdivided into 14 sublevels. These sublevels offer a more fine-grained level of cognitive difficulty compared to Bloom. In our recasting of the questions from Bloom to Marzano we could cover 8 of the 14 Marzano subcategories. These more fine-grained levels offer hope that we might be able to more accurately determine which types of cluster behaviours correspond to more specific tasks as defined within Marzano. We will use the combined data from experiments 1 and 2 to enhance the sample size and improve the quality of the analysis.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.19396	0.15449	0.13967
60,30,10	0.26245*	-	0.18749	0.17548
70,20,10	0.3361*	0.07365	-	0.13053
80,10,10	0.5412*	0.27879*	0.20514*	-

Table 5.11 Tukey-Kramer Analysis Marzano Level 1 Sublevel 2 for Experiment 1 and 2

When we recast our Bloom levels, all three sublevels within Marzano's level 1 are represented. As can be seen in Table 5.11, Marzano level 1 sublevel 2 (Marzano 1:2) shows there are significant differences between the clusters. Table 5.10 shows that significant differences exist at Marzano 1:1. However, they only seem to be significantly different for the 50,30,20 cluster. We hypothesize that if the number of participants is increased, we should see the same significant differences between all the clusters.

	A50,30,20	**A60,30,10**	**A70,20,10**	**A80,10,10**
A50,30,20	-	0.24477	0.32048	0.21365
A60,30,10	0.31055*	-	0.3583	0.26706
A70,20,10	0.4267*	0.11613	-	0.33781
A80,10,10	0.5548*	0.24429	0.12817	-

Table 5.12 Tukey-Kramer Analysis Marzano Level 1 Sublevel 1 for Experiment 1 and 2

We had hoped the more fine-grained level of cognitive difficulty would provide a more fine-grained grade prediction such as a C+ or A-. Unfortunately, the clusters were only able to predict the same letter grade differences that we have been able to predict at the coarse-grain level of cognitive difficulty that we found in experiment 1 and in the more coarse-grained analysis performed above in experiment 2. The extra levels of granularity provided by Marzano do not improve the predictability.

5.4 Summary of Results for Experiment 2

There were two major results found in experiment 2. First was the confirmation of our results from experiment 1. Using Bloom's Taxonomy combined with the centroids found in the first experiment, we were able to both predict grades and show the same significant differences between the clusters as we found in experiment 1. Furthermore, the use of Marzano's Taxonomy at a coarse-grained level was able to provide the same levels of grade prediction as was found using Bloom's Taxonomy.

Second, we found that a more fine-grained level of cognitive difficulty as found within Marzano's sublevels could not allow the prediction of finer-grained grades compared to the more coarse-grained levels of both Marzano and Bloom's Taxonomies. This suggests that Marzano does not add much beyond what Bloom has already provided in helping analyze student problem solving behaviour when solving reading comprehension tasks, but the work is by no means definitive on this matter. There is more research that could be done to see if there are other behaviours that can be predicted with a more fine-grained level of cognitive difficulty.

The next chapter will directly address the issue of answering a Bloom level 6 question within the context of an ill-defined domain. It will also contain a few low-level Bloom questions to add further confirmation to the results from experiments 1 and 2. We will not be pursing Marzano's Taxonomy further at this time as we wish to focus our efforts on examining the interaction of ill-defined domains with the clusters we have discovered. This means that we will ask a question based on content that falls into the ill-defined domain and see how well our clusters can identify student success as determined by grades.

CHAPTER 6
EXPERIMENT 3

In this third experiment, we are hoping to be able to examine several questions with the overarching goal of firmly establishing the role of educational taxonomies within advanced learning systems. First, as before, we want to explore student interaction data as it relates to a student answering a high level Bloom question, but this time using a different context than we have in our previous experiments. Our goal in this experiment is not only to further confirm our findings from previous experiments but also to investigate Bloom Level 6 in more detail. Our previous experiments used high level Bloom questions; however, the answers we expected were only a few paragraphs. This third experiment is designed to elicit an answer from the student that is at least one and a half pages in length and contains at least two citations from the documents provided. This is to elicit more synthesis from the student with respect to the material they are presented with. We are interested to see if the need for the student to produce longer answers affects how a student interacts with the documents.

Second, we want to pose a question that clearly can be categorized as the kind that comes up in ill-defined domains. The goal is to add to the evidence already provided in the first two experiments that the methodologies we are proposing will work within an ill-defined domain.

Third, we would like to explore the role of keystroke data captured as the students typed the answers to the questions into the system. Although this data was captured in the previous experiments, it had not been used in the experimental analysis in order to focus on reading behaviour, rather than writing behaviour.

There are many other questions that we could ask in this experiment. For example, is there an interaction between different types of documents and how a student reads, scans and scrolls through these different types of documents? Does the amount of time that a student has to answer a question change how they interact with the document? These are new questions that have been opened up, but we will not be seeking to answer these directly in Experiment 3.

6.1 The Study

The third experiment took place in the 2014-2015 academic year at Saskatchewan Polytechnic in Saskatoon within the adult grade 12 English program. Three separate sections of the Grade 12 English A30 course taught by two different teachers were used: two sections in the first semester and one section in the second semester. The students in these classes were required as part of their workload to do an assignment that involved reading documents (on cultural diversity within Canada) and answering questions about them, and were given the option of using the same interface as we developed for the first two experiments (although they could use pencil and paper, in which case their data was not included in my experimental analysis). The teachers for each of the participating classes graded the answers for the students who were enrolled in their courses. The teachers used the same rubric in each course, which is handed out to the students at the beginning of the assignment, to assess the students. A prize of one tablet was awarded to a student in each class that participated, with the winner being chosen by a random draw consisting of those students who participated.

The experiment was divided into two parts, experiment 3A and 3B, run with different students: 3A were the students in the one first semester class; 3B were the students in two second semester classes. For both parts, the following took place:

- The students carried out the experimental task as a part of their regular course work and their grades for this assignment formed part of their final grade within the course. The teachers marked their respective students to maintain course and grade continuity for their students. The students were provided a rubric from each of their teachers outlining the expectations and grades for their work prior to the experiment beginning. The teachers had gotten together to develop one rubric that was used by both of them to grade the students.
- All assignments were graded by the student's respective teacher according to the rubric.
- The students were allowed to opt out of the experiment and perform the same assignment as a paper based assignment rather than using the computer interface provided by our experiment.
- There were four questions being asked of the students: three questions at Bloom levels 1 and 2 and one journal response question at Bloom level 6. The low level Bloom

questions test the robustness of the results already seen in earlier experiments. The high level Bloom question allows us the opportunity to see how well the approach works within typical difficult ill-defined domain problems.

- The journal response question required the student to write at least one and a half pages and involved incorporating information from multiple documents to answer the question. The length of the answer is deemed by the teachers to be appropriate for the type of students answering this type of question.
- The types of documents provided to the students included poetry (two documents), quotations, personal journal, and articles (two articles). The students were required to reference two quotes from two different articles that applied to their response. These documents were the same for both experiment 3A and 3B.

This experiment consisted of four questions. The first question was a Bloom level 6 question, clearly an open ended question of a type typically posed in an ill-defined domain:

> *Q1) Write a journal response following the advice discussed in class. Be sure to write enough, make sure you include ideas from the readings you are given, and include higher levels of thinking in your response. Do NOT summarize the texts, although you may include small bits of information that help explain the points you are making in the journal response. Here is the question. Comment on the following grand narrative that many people worldwide believe about Canadian identity: "Canada is a democratic, multicultural country free from racism and violence. Canadian citizens are caring and tolerant people that have a global reputation for peacekeeping." This is one of the quotes that Sheelah Mclean likes to discuss with audiences around the world. For those who don't know Sheelah Mclean, she is one of the founders of Idle No More, a group that stands up for Aboriginal rights in Canada. Do you agree or disagree with this quote (grand narrative)? Why or why not? Be sure to include evidence from the readings that have been given you.*

As can be seen by the question, there is no specific right or wrong answer. The answer to this question is based upon previous personal experience, class discussions and opinion being filtered through the lens of the documents that are read. The students are to point out various sections of text that they have read and provide arguments to justify their particular perspective using these references. This type of question is one of the most difficult for an automated system to mark [98].

The teachers were present to answer students' questions about the requirements of the assignments and to ensure that they were providing the type of answers the teacher was looking for. For Experiment 3A both the teacher and I were present for the entire three hours. For Experiment 3B the teacher and I were only present for the portions of the experiment where the students were provided in-class time to work on the assignment over the two-week period. In both 3A and 3B the students were provided instructions and options to continue their work from home if they felt that it was necessary.

6.2 The Participants

Over the three experimental groups, there were 78 participants from Saskatchewan Polytechnic. The participants in both Experiment 3A and 3B were adult students who are returning to school to obtain their grade 12 courses prior to moving on to post-secondary education. For Experiment 3A, 21 participants were used and the remaining 57 were in Experiment 3B. The students were from different classes and in different semesters. The background of the students ranged from those who dropped out of Canadian high schools for various reasons to students from other countries with degrees looking forward to upgrading their education to be accepted into Canadian post-secondary education.

6.2.1 Instructions

The students were given a simple set of instructions along with a short overview of the interface. The instructions were that they were to read the documents provided an answer the questions that are based upon the documents they just read. The overview of the interface involved ensuring that the students knew where the documents were located, where the questions were located, where the text of the document would appear and where they could type in their answer and then to press the submit button when they were finished the experiment. The students were told that the system would record how they read and how they answered the questions (see Appendix 3 for full consent form). They were not given any instructions on interaction methods they should use to complete the experiment. In keeping with trace methodology approaches [4], all the interactions with the content and questions were recorded and time-stamped. These include events such as mouse click, mouse wheel, which item was clicked or selected and so on. The students in 3B were shown how to login to the system from

home if they wished to work from home on the assignment. They were told that they would be given class time during the two week period to work on the assignment.

6.3 The Procedure

The interface that the students used is the same interface used in Experiment 1B and has been previously described. As in the other experiments all of the data collected were time-stamped to the nearest millisecond and the resulting data captured were the same interface level keystrokes that were captured in the first two experiments.

For Experiment 3A, students from one section of the adult grade twelve English A30 course participated. The class schedule was rearranged to allow for a single 3 hour period to occur. The students were given the single 3 hour period to work on the assignment. If the students were not able to complete the assignment in the time allotted, they were given a further day to complete the assignment as homework at home. They were not given any further class-time to work on the assignment.

For Experiment 3B, students from two adult grade twelve English A30 courses participated. The sections were run concurrently but with different teachers and at different times of the day. The course content being covered was the same for both sections. There were two different teachers teaching the sections. The experiment ran over a two week period of time with the same questions, rubric, and documents as in experiment 3A. All of the participants completed the assignment within this time frame. The students in Experiment 3B were given portions of multiple class periods to work on the assignment over the two week period. The students could log into the system and work from home if they desired.

6.4 Results

There were a total of 78 participants who generated 248,280 events for the four questions. For experiment 3A several of the participants elected not to complete the lower level Bloom questions in favor of working only on the long answer question. They were not included in the results. Since time was a constraint in Experiment 3A, this was expected.

6.4.1 Results for Experiment 3A

Table 6.1 shows the results for the ANOVA for Experiment 3A. We chose to show 3A since its experimental design, and in particular, the time constraint was very similar to both experiments 1 and 2. Since Experiment 3B changes the time requirements for how the experiment is done, we did not feel that this would make a fair comparison even though its values are also similar to the table below. Experiment 3B results will be discussed in terms of the Tukey-Kramer analysis later on. The ANOVA results from experiment 3A are very similar to the results from experiments 1 and 2 (shown in Tables 4.1 and 5.1).

Bloom Level	**F**	**P**	**F-Critical**
1	*39.49	9.21E-12	2.85
2	*11.907	.0003	3.28
6	*51.42	5.24E-11	3.27

Table 6.1 ANOVA for Bloom Levels for Experiment 3A

6.4.1.1 Grade Prediction by Cluster Type for Experiment 3

Again, we grouped each of the questions into their corresponding Bloom level and mapped each of the grades to each of the clusters. Table 6.2 shows the mapping between the various clusters for each of the Bloom levels and the grades that are associated with them. Table 6.3 shows a more in depth look at Bloom Level 1 Medium Light Reading from Table 6.2. We can see as happened in other experiments that we could increase the % accuracy of cluster inclusion if we took a deeper look at what was happening within this cluster over the portions of the answers.

One of the major issues with this experiment was that we could not get the level of predictability closer to 100% for the Heavy Reading Bloom Level 6 as seen on Table 6.2. All but one of the participants ended up in this category. Although the cluster is able to predict the correct grade almost 80% of the time, the positional analysis was unable to arrive at a suitable solution because the answers varied from individual to individual. The ill-defined nature of this type of problem did not allow for us to find uniform specific locations of text where we could perform the positional analysis like we have done in Table 6.3.

Cluster	Bloom	Grade	% Accurately Predicted
Light Reading	1	A, F	(Grade = A) 60%
Medium Light Reading	1	A, F	(Grade = A) 50%
Medium Heavy Reading	1	A, F	(Grade = A) 80%
Heavy Reading	1	A, F	(Grade = A) 53%
Light Reading	2	A, F	(Grade = F) 75%
Medium Light Reading	2	D	(Grade = D) 100%
Medium Heavy Reading	2	A, F	(Grade = A) 50%
Heavy Reading	2	A,C,D, F	(Grade <= D) 63%
Light Reading	6	N/A	N/A
Medium Light Reading	6	N/A	N/A
Medium Heavy Reading	6	A	(Grade = A) 100% n = 1
Heavy Reading	6	A,B,C,	(Grade >=B) 79%

Table 6.2 Experiment 3 Grade Prediction by Cluster Type

Cluster	Bloom	Grade	% Accurately Predicted
Light Reading	1	A	100%
Heavy Reading	1	F	100%

Table 6.3 Positional Analysis of Bloom level 1 Medium Light Reading

6.4.1.2 Low Level Bloom Analysis for Experiment 3A

Table 6.4 shows the Tukey-Kramer analysis for the Bloom Level 1 questions. It is the lower level Bloom questions that provide further confirmation of the findings from the first two experiments. The high level Bloom question, which is analyzed later, did not produce the same results we found in the first two experiments. The low level Bloom questions showed that the participants tended to group in the 50,30,20 (Light Reading) and the 80,10,10 (Heavy Reading) clusters. It appears from Table 6.2 that the participants either scored well or outright failed for the Bloom Level 1. Table 6.3 shows that when we examined what occurred directly over the answer that a slow heavy reading method was not able to recall the specific answer compared to scrolling to where the answer was located to answer the question. The 60,30,10 cluster does not

appear to be significantly different from the 70,20,10 and the 50,30,20 clusters. The GCI analysis should shed some light on this.

The analysis for the Bloom Level 2 questions are similar to the analysis performed for Bloom level one. The statistics can be found within Appendix 3.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0.1914	0.14914	0.10851
60,30,10	0.14757	-	0.20779	0.18085
70,20,10	0.23704*	0.08947	-	0.13534
80,10,10	0.4301*	0.28255*	0.19308*	-

Table 6.4 Tukey-Kramer for Bloom Level 1 Experiment 3A

If we look at Figure 6.1, the Gabriel Comparison Interval indicates that the 60,30,10 and 70,20,10 clusters have a larger variance compared to the other two clusters. This is due to the low value of N (n = 2) for both clusters. It seems that the participants choose to either perform Light Reading (50,30,20) or Heavy Reading (80,10,10). This makes sense given the type of assignment. In order to find a specific low level answer a light reading strategy where more scrolling and scanning is beneficial. However, the first question is the long difficult question which should cause a Heavy Reading strategy so that the participants can understand and cognitively integrate the material. Since the first question in the list is the long answer question, this may have unintentionally promoted the idea of Heavy Reading for all of the questions. For those students who answered question 1 first and then moved on to answer the remaining questions; we expected to see more scrolling and scanning to locate the answers for the other questions they had previously read while answering question 1. However, for those students who took the opportunity to answer the easier questions first so as to allow these questions to inform the first question we expected to see more Light Reading first. These results seem to indicate that both these were the strategies that were used by those that answered these questions. However, as in the previous experiments it was those students who performed the Light Reading (50,30,20) who received the higher grades on the lower level Bloom questions.

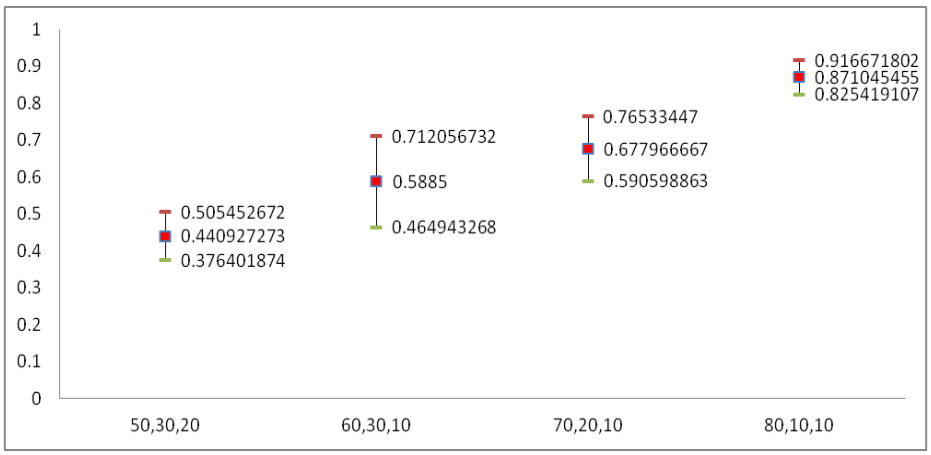


Figure 6.1 Means with Gabriel Comparison Interval for Bloom Level 1 for Experiment 3A

6.4.1.2 High Level Bloom Analysis for Experiment 3A

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-	0	0	0
60,30,10	0	-	0	0
70,20,10	0	0	-	0
80,10,10	0	0	0.22949*	-

Table 6.5 Tukey-Kramer for Bloom Level 6 Question Experiment 3A

Table 6.5 show the Tukey-Kramer analysis at the Bloom level 6 question from experiment 3A. We see from Table 6.5 that there are only two clusters that become evident. If we look at the N of the 70,20,10 cluster, there is only one student who fell into that cluster. The remaining students all fall into the 80,10,10 cluster. This suggests that the question we posed seems to have a direct impact on the strategy taken by the students with the three hour time constraint – essentially Heavy Reading was required. This also did not allow us to predict grades since the majority of the students all fell into one cluster. This was not a result we have previously encountered in the other experiments.

6.4.2 Bloom Level 6 Analysis for both Experiment 3A and 3B

For Table 6.6 the corresponding ANOVA was not included in this analysis since it did not add new information other than report that significant differences were found and will be explained by the subsequent Tukey-Kramer analysis. The ANOVA can be found in Appendix 3.

Table 6.4 shows the Tukey-Kramer analysis at the Bloom level 6 question for students in both experiments 3A and 3B (with N = 78). The results in Table 6.6 show that there were really only two major categories of clusters utilized for the first question: one strategy where more scanning and scrolling (Light Reading Cluster) occurred and one where more Heavy Reading occurred. We have seen similar results in our earlier experiments. Since the students needed to perform a lot of reading in order to successfully answer this question, we did expect to see a lot of heavy reading. We did not expect to see the large amount of scanning and scrolling as indicated by inclusion in the Light Reading cluster. However, it should be noted that the students found in the Light Reading (50:30:20) cluster performed poorly (no grade higher than a D).

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-		0.11899	0.06567
60,30,10				
70,20,10	0.15826*		-	0.11542
80,10,10	0.3491*		0.19085*	-

Table 6.6 Tukey-Kramer for Bloom Level 6 Question for Experiment 3A and 3B

6.4.3 Analysis from Student Answers

Since our Heavy Reading (80,10,10) cluster ended up containing the majority of the participants we needed to take a look to see if there were other factors that might allow us to differentiate between the various grades awarded to the students in this cluster. To this end, we examined the answers that were provided by the students to see if this could help us locate patterns in the student's behaviour that we could use to predict grades.

Within the Heavy Reading Cluster, there is variance around how much reading, scanning, and scrolling take place. So within this cluster the variance can range between participants who did more reading (98,1.6,0.4) compared to other students who did more scanning (77,17,6). In order to see if we could differentiate between the grades we looked to see if the amount of

variance in the amount of scanning and scrolling within the Heavy Reading clusters would provide a better predictor. To that end, we analyzed the answers the students gave with the type of scanning and scrolling they exhibited within the Heavy Reading Cluster.

What we saw was two distinctive patterns that emerged from this cluster on how the students used the documents to justify their answers. The first pattern was where a student wrote their response to the question and then sought out justification from the documents to back their argument. This was observed in the data where we would see a lot of keyboard activity taking place followed by a lot of scanning and scrolling. This was not a strategy we had anticipated for this type of question. Some of these students received good grades because they could find and incorporate information from the documents they read into their answers. This turned out to not be the best predictor since it did not always predict a good grade. There were those who did not locate information from the documents they read and so were not able to properly incorporate quotes to support their arguments and consequently received poor grades. This was often indicated in their answers. They would perform some typing, do some scanning, and then continue answering but not answering anything related to their positions within the document.

The second strategy was to thoroughly read the documents and then form an answer based upon the student's own experiences and what they have read. Again, those students who incorporated the information from the documents into their answers received good grades. Conversely, there were those students who read the documents yet, for some reason, did not incorporate information from the documents into their answer. These individuals received a poor grade.

This meant that for both strategies we were not able to reliably predict any grades by inclusion in clusters. It appears that for Bloom level 6 questions of any depth our clusters no longer can be used to predict grades with any reliability. In the next section, we examine another source of data with the hope of being able to more accurately predict grades.

6.4.4 Keystroke Data

The next analysis we performed was to look at the writing data as the students typed their answers. The writing data contains the duration of time between successive keystrokes by the user. Since writing out an answer often results in a short duration of time between keystrokes, a

cutoff time could be used to determine when an individual is typing. For a fast typist, the duration between keystrokes will be milliseconds. For a more inexperienced typist, the time between keystrokes would be longer. Given the range of typing we found in the data, a time of less than 2 seconds was deemed appropriate to rate a classification of "**typing**" while using our interface.

There were other longer durations that were found between keystroke events in the data. When one thinks of what is occurring while you are typing there are a couple of activities that come to mind. The first is that there is a shorter pause ("Composing") between keystrokes as you think about the next word or set of words that you are going to type. Second there are longer pauses ("Thinking") when you are actually working on the next major thought that you would like to type. Lastly, there are very long pauses that may be either classified as user distraction or where you are composing a paragraph or section that you would like to write. We impose an arbitrary time limit of one minute as being the cutoff for thinking. These cutoffs are a first approximation for writing categories, drawn by analogy from the reading cutoffs, and we became unsure of what might be happening beyond one minute. There may be interesting pedagogical reasons for pauses greater than 1 minute, but we did not pursue these rare instances. This is not to say that thinking does not occur beyond this point, but that other factors such as the participant being distracted increase in probability the longer there is no activity. Keep in mind that this type of analysis would not work for just any arbitrary keyboarding situation. However, when you have a specific educational task that is being done, the participants are on task and they are working on a specific answer, so we are able to make some of the above assessments.

Composing was defined as pauses that are longer than 2 seconds but less than 5 seconds, suggesting that the student is searching for the next word to type. **Thinking** would be mapped as pauses greater than 5 seconds but less than a minute, where students are presumably coming up and organizing their ideas. This leaves us with the following cutoff times for pauses between keystrokes:

- Typing < 2 seconds
- Composing <= 5 seconds => 2 seconds
- Thinking > 5 seconds < 1 minute

Within the data there were 189,643 keystroke events captured only 321 keystroke events had a duration greater than 1 minute between successive keystrokes. Furthermore, only 4 instances occurred where there was slightly more than 10 minutes between keystrokes with the longest pause being 12 minutes. There were only 19 instances where there was between 5 to 10 minutes between keystrokes out of the 189,643 keystrokes captured. This seems to indicate that the participants were on task the majority of the time they were working on a problem and that our one minute cutoff is a reasonable value that does not exclude much data.

6.4.5 Clustering Keystroke Data with Time Cutoffs Using Hard Coded Clusters

One cannot help but notice the similarity to the reading comprehension time cutoffs that we used in our experiments. Within education, reading, writing and arithmetic are called the 3 Rs. Reading and writing have a long history of being linked together and justifiably so [99]. During the reading process, the brain requires time to perform various activities to allow for the comprehension process to happen. Our experiments seem to validate this thought as the timing between the various reading events correlates well to grades. Similarly, the process of writing answers and the timing between keystrokes appears to have similarities to our results from reading. Due to the similarities between reading and writing we decided to see if the typing times would cluster in a similar manner to the reading comprehension data, we clustered the data using the hard-coded centroids we had discovered in experiment 1.

Using the hard coded centroids we had used for the reading comprehension clusters, we ran the clustering algorithm to see how the clustering algorithm would work on the keyboard data. We ran an ANOVA (found in Appendix 3) on the results which returned statistically significant differences for the clusters representing the pause time between keystrokes. Since in the past, we have found most of our statistical significance with the larger data sets we chose to put the data for 3A and 3B together first rather than work on a smaller subset of the data. Table 6.7 show the Tukey-Kramer analysis for the keystroke data for the Bloom Level 6 question for experiments 3A and 3B combined. As can be seen by the analysis only three of the clusters are represented. An individual belonging to the 50,30,20 (Fast Typing) cluster would spend approximately 20% of their time just performing rapid keystrokes (Typing). The majority of the pauses between Typing were in the millisecond range providing strong indications that they were just typing. They would also spend 30% of their time pausing between 2 to 5 seconds between

keystrokes (Composing) and approximately 50% of their time pausing for more than 5 seconds between keystrokes (Thinking). Individuals belonging to the 80:10:10 (Heavy Thinkers) would spend approximately 80% of their time thinking with only approximately 10% of their time devoted to typing and composing. The 70,20,10 (Medium Heavy Thinkers) would spend approximately 70% of their time thinking, 20% of their time composing, and approximately 10% of their time typing.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-		0.10352	0.05758
60,30,10				
70,20,10	0.14037*		-	0.10007
80,10,10	0.3314*		0.19098*	-

Table 6.7 Tukey-Kramer Keystroke Analysis for Bloom Level 6 Experiment 3A and 3B

Next we analyzed if inclusion in a specific cluster was indicative of grade. Each of the clusters represented contained both high and low grades within the clusters. However, if we looked at the 3 ratios (Thinking, Composing, Typing) that were used to define which cluster they were in, the "Composing" (between 2 to 5 seconds) had cutoff values that separated the grades within the cluster for the majority of the cases. Composing between keystrokes seems to imply that there is some sort of mental activity taking place that differentiates the grades that a student gets. So we see that the cluster the student ended up in indicates broad performance bounds with respect to keystroke behaviour. For example, people in the 80,10,10 (Thinking) category spent approximately 80% of their time typing with longer pauses between keystrokes for thinking to occur, while the people in the 50,30,10 (Fast Typing) category spent a 30% of their typing time with short pauses. However, out of the three ratios, Composing seemed to provide a cutoff that was predictive of grades. This is demonstrated in Table 6.9 and will be discussed later. Again as in the previous experiments, the prediction of grade is relegated to being able to predict a letter grade.

6.4.6 Effectiveness of Hard Coded Centroids

The use of the hard coded centroids makes some intuitive sense. However, in practice this does not bear out with the data. What we see is that there are only three clusters that show up. This means that one entire cluster from the hard coded centroids is not even represented.

Furthermore, Figure 6.2 shows the 70,20,10 cluster with a large confidence interval which indicates that the N = 6 is not large. The majority of the participants end up in either of the extreme clusters. Although the 70,20,10 cluster has a small N, it does make a difference with respect to some participants in aiding in the prediction of the grades as seen in Table 6.7.

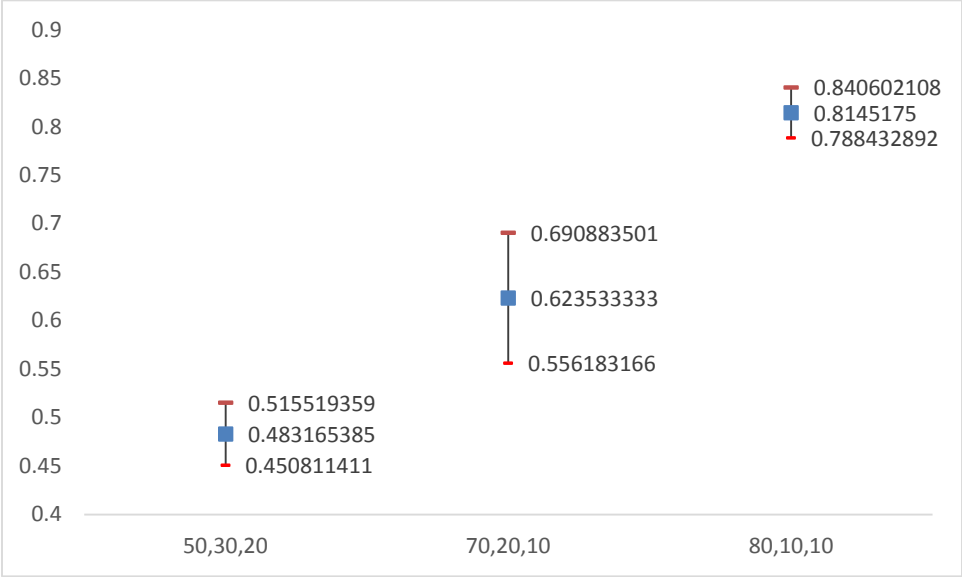


Figure 6.2 Means with Gabriel Comparison Interval for Bloom Level 6 for 3A and 3B for writing data

6.4.7 Results Hold for Smaller Subset of Data

Given the time constraints on 3A, we felt it would be interesting to do a separate analysis of the keystroke clustering on just the 3A data, since the clusters themselves were most pronounced in our reading analysis when the students had to work under time constraints. The sample size for the 3A data set is 1/3 the size of the combined data set. If our results hold with this smaller subset of the data that is more closely related to our other experiments, it stands to reason that these results will also hold for both our first and second experiments.

	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**
50,30,20	-		0.12142	0.0761
60,30,10				
70,20,10	0.13105*		-	0.12214
80,10,10	0.31464*		0.18359*	-

Table 6.8 Tukey-Kramer Keystroke Analysis for Bloom Level 6 Experiment 3A

Table 6.8 show the Tukey-Kramer analysis for just experiment 3A (corresponding ANOVA in appendix 3). Not only did the same clusters appear in the combined 3A and 3B analysis of writing, but they also appeared with similar significant differences. Furthermore, the cutoffs we found in the combined analysis also held in 3A as well.

6.4.8 Keystroke and Reading Ratio Cutoff Values for Predicting Grades

As mentioned previously the clusters for the keystroke data were able to predict the majority of the grades that were found in the clusters. However, they were not able to predict all the grades found by the clusters.

Keystroke Cluster Type	**Composing Ratio (2 to 5 seconds)**	**Reading Ratio**	**Grades**
50,30,20	Ratio > 0.10	Ratio < 0.75	Grade < C
70,20,10	Ratio > 0.07	Ratio < 0.80	B
80,10,10	Ratio > 0.047	Ratio < 0.91	A

Table 6.9 Combination Short Pause Ratio and Reading Ratio for Grade Prediction

Table 6.9 shows that a combination of both Reading Ratio cutoffs and Composing Ratio cutoffs for each of the clusters provides the best predictor of grade. The calculation for the Composing Ratio is a composition of all the time that the student spent writing their answer and dividing it by the amount of time between keystroke events for that level of difficulty. For example, if a participant belonged to the 50,30,20 (Fast Typing) cluster and had a composing ratio of greater and 0.10 and had a reading ratio of less than 0.75 they received a letter grade of C at best. This classification system was able to correctly classify all but 1 case.

6.5 Summary of Results

There were some interesting results from this experiment 3:

- First, at the lower Bloom levels, it was able to provide further confirmation of the first two experiments.
- Second, at the higher levels of Bloom, our current set of clustering was not able to effectively determine grades as was done in the first two experiments. This is probably due to a few factors. The type of problem was one that required a much longer answer

than had previously been attempted. The time constraints of 3A meant that the majority of the participants only performed the Heavy Reading strategy. This left us with no method to differentiate between grades. Experiment 3B did not have the same time constraints that we found in 3A and as a result we found some students opted for a Light Reading strategy which resulted in both poor and good grades. This seems to be counter intuitive since the students have more time to read the material we would expect everyone to use a Heavy Reading strategy. Our guess as to one possibility is that given the students felt they had lots of time; they procrastinated until the last minute and then were in a rush to complete the assignment. If this was the case, this would explain the poor grades. For those students in the Light Reading cluster that received good grades, the quicker movement through the documents may have provided the students with the general overall feel and they were able to take those concepts and weave them effectively with their own insights and then find and locate the supporting documentation within the articles to support their insights. Irrespective of the reason, these grades for the Light Reading and Heavy Reading clusters indicate that there are cases where the results from the first two experiments do not hold. However, we were able to find new methods of overcoming this problem.

- Third, the addition of keystroke data analysis to the current reading analysis provided a method to account for all but one grade. Although the prediction capability was only at the letter grade level but it is the same level of prediction as with our other experiments.
- Fourth, we were able to effectively predict grades on a problem that most advanced learning systems would not be able to handle.
- Fifth, we opened a new area of research that needs to examine what exactly is happening during the pauses between the keystrokes when students are writing. We know that there are some similarities to the reading comprehension results in our three experiments, but we are not sure about what processes are going on during the writing portion of the student answering a question.

CHAPTER 7
CONCLUSIONS

7.1 Overview of the Three Experiments

The first set of experiments discovered the metrics that helped measure the types of reading that can be done, at least, in the context of students answering questions about documents. The factor that we were most interested in was the level of cognitive difficulty of the question that the student was currently working on. Our results showed that there were certain types of strategies that were successfully employed when the problem had a low level of cognitive difficulty and there were different strategies that were successful for the higher levels of cognitive difficulty as categorized by Bloom's Taxonomy. The reading, scanning and scrolling ratio was calculated for all the content that was read while the student was working on a particular question. Additionally, there were indications that the level of granularity, with respect to document position where the reading, scanning, scrolling ratios were applied, had an impact on the predictability of the grade for a given question. When we looked at the reading, scanning and scrolling ratio over the portion(s) of the document(s) that contained the answer to the question being currently worked on, our predictive accuracy increased slightly. By increased predictive accuracy, we mean that we were able to account for all the grades for that question.

The students did not always choose the same reading strategy for each document when there was more than one document. For example, we saw some students use the same reading strategy for all of the documents (Heavy Reading) but they did not complete the experiment as there was more to do than they had time for. Students who changed their reading strategies between documents and even within documents were more successful. For example, a student might start heavily reading a document only to find that it does not seem to be necessary for answering the question they are currently working on. Once this realization occurs, they switch strategies to more of a scanning strategy for the remainder of the document, unless they come across a portion of the document that is relevant to the answer at which time they again change reading strategies.

The second set of experiments was aimed at first confirming the results of the first experiment and second, to determine if categorization of the questions at a more fine-grained level compared to Bloom's Taxonomy would increase the reliability and predictability of our

results. To this end we reclassified both the first and second experiment's questions using a competing classification hierarchy developed by Marzano. With respect to our first aim, we found that the clusters allowed the same predictions to be made in the second experiment as in the first, thus providing strong encouragement that we did not over fit the data in the first experiment. For the second aim of experiment 2 we were able to reclassify all of our questions from both experiment 1 and experiment 2 into Marzano's Taxonomy, but we did not achieve any improvement in predictability of grades. Since in our experiments, both taxonomies provided predictive capabilities using the same clusters of reading, scanning and scrolling ratios, we do feel that this experiment shows that both taxonomies are addressing the same set of underlying cognitive properties.

Experiment 3 confirmed the first two experiments with respect to predicting performance at the lower levels of Bloom. The Bloom level 6 question used in experiment 3, however, showed the limit to the usefulness of the clusters discovered in the earlier experiments. The context for this type of question was markedly different from those of our previous experiments. The length of the answer was dramatically longer, and the open-ended nature of the question would be a challenge for any ALS to predict grades. When the reading clusters discovered earlier did not yield any interesting connections to a student's performance, we turned to analyzing the student's writing, in particular, analyzing the pauses between keystrokes, to see if anything interesting would show up. While biometrics has made use of keystroke timings for identification purposes [100], and lexical analysis techniques such as LSA have been used within automatic grading systems to determine if an answer is correct [101], making predictions of grades based on the length of pause between keystrokes when writing is a unique contribution. However, the initial results from this analysis require a more specifically designed experiment to confirm and extend the approach. In particular, we will need to try to determine what is actually occurring during these pauses and determines if the thresholds are meaningful in terms of student cognitive activity. Does a pause of a shorter duration indicate that the student is forming the next word or sentence? Are the longer pauses indicative of the student thinking of the next sentence or paragraph? These are some of the questions that need to be addressed in future experiments.

7.1.1 Limitations of the Experimental Results

The experimental results, future work, and conclusions should be interpreted in the light of some of the limitations that are found within our work. Our sample sizes for experiments one and two were both small. This limited the number of participants who tested each of the different levels of Bloom's / Marzano's Taxonomies for the experiments. This forced us to keep our interface and experimental design extremely similar so we could group data across multiple experiments and draw more robust conclusions. A larger number of participants in a single experiment that tested all the Bloom levels multiple times within the same experiment would, of course, further add to the robustness of our results and confirm the validity of our approach of drawing conclusions across multiple experiments.

Our interface design was another limitation of this set of experiments. The reading window was kept small in order to provide us with a reading speed as well as the position within the document that was currently being read. In most Learning Content Management Systems (LCMSs) the size of the reading window is much larger. Since we wanted to be able to track participants who might work from a computer at home or any computer within the school, we had to limit the size of the reading window. Eye tracking software would allow for a much larger and more natural interface to be utilized; however, this would limit the number of participants who could simultaneously run the experiment due to the specialized hardware and software that is required in order to accurately perform eye tracking. Interestingly, eye tracking software would allow for experimentation on image/diagram comprehension as well as aiding in determining if a participant is distracted. This would involve creating a dwell time ratio over the image/diagram to determine which parts of the image the participant is concentrating on. The camera could also be used to help determine the affective states of the participant during the experiment.

Our experimental design allowed for unconstrained movement of the participant between documents, questions and the answers they were working on. In many courses built in a LCMS, this level of flexibility is possible, but it is rarely utilized. It is this freedom of interaction with the interface that our ratios captured and these results may not transfer too more constrained environments. There are times when a student is constrained to being able to answer the question without having the ability to look back and refer to content. Will the reading, scanning

and scrolling ratios still hold when a student cannot interact with the documents as they work on answering a question?

7.2 Conclusions

The contribution of this work should have its most impact in ill-defined domains. Reading is a necessary component in so many such domains, and a tool that analyzes reading behaviour to make predictions about grades a student receives on questions can be a benefit to many different systems and tools. Of course, reading comprehension is so much more than simply a score for a question. Nevertheless, asking and grading questions are one of the primary methods used to determine if a student understands the content they have read. In a classroom setting a teacher can pick up various cues that a student might present to help the teacher make the assessment that a student either understands the content presented or that they might not understand the content that the student has read. In an online environment, an automated system does not have the ability to perform as a human. Despite the advances made in the advanced learning technology, we cannot yet replicate the flexibility of a teacher who can teach multiple subjects and respond effectively to a student's need in a diverse set of domains. This work provides a framework for a tool that holds the promise to be generalizable across domains. In section 7.3, I discuss the various ways that an ALS might make use of a diagnostic capability such as that demonstrated.

This work has demonstrated that Bloom's Taxonomy (as well as Marzano's) provides a framework with which to help find patterns that predict student success or failure. In fact, this framework is necessary, for without such a framework, no useful patterns have been found: in short, the level of cognitive difficulty of the question to be answered must be known to make a prediction of how well the student will answer it. The 4 main clusters we have found have been shown to be resilient across the domains we have tested them in. The clusters have also been shown to work across various educational levels of students, from high school to college to graduate studies. In fact, others have started to build on this work. In [102] learning progressions are used in a similar vein to both Bloom's and Marzano's Taxonomy as they relate to levels of cognitive difficulty.

The three Rs have been foundational in so much of the educational system. It should not really come as a surprise to find that both the reading and the writing components of student interaction data seem to be tied closely together. Our last experiment demonstrated that both reading and writing were necessary in order to make letter grade predictions on how well a student preforms on a high level Bloom problem. Furthermore, the open-ended nature of the high level Bloom question in Experiment 3 provided both a challenge to our current clustering model and an opportunity to explore the new territory.

Analyzing the writing of students as they answer questions (not just their reading behaviour), in particular examining the pauses between the keystroke data, is an area where future work will need to be done. To the best of my knowledge, no one has yet tied the pauses between keystrokes to the performance of a student, as predicted by grades, working on an educational problem at a particular level of cognitive difficulty. We have noticed that the pauses in keystrokes do not have the same number of clusters as in the reading clusters. We discovered that the composing ratio is predictive of grades when combined with reading ratios. We need to first, confirm the results of this experiment and then move on to try to discover what mental processes are occurring during the keystroke pauses. The questions about what exactly is happening during the pause also makes for an interesting research investigation: are the participants, thinking, composing, looking for the next key to press? Are these pauses equally predictive within other contexts and domains at Bloom level 6? Do these pauses hold for the lower levels of Bloom? Are the hard coded clusters we used the only or the best for making predictions? Are there other different behaviours that we could possibly cluster? All these questions point to a new research area that needs to be explored.

7.3 What other studies could this data provide?

Throughout the three experiments, we have primarily focused on the clustering of the reading ratios and their predictability when contextualized by the level of difficulty in Bloom's Taxonomy of the task the students are working on. However, there are other aspects of the data that we could analyze. One such aspect is to take into account how the student moves through the reading, scanning, and scrolling activities. In our experiment, we create "buckets" (clusters) of the ratios that exist as the student works through the various questions and documents. There

is a sequential aspect that exists within the data that we did not examine or fully explore but was noted as we looked through the data. For example, a student might perform reading first, then maybe rereads a portion of the document before moving on. Are there predictive patterns that exist within the sequential data that would help predict student success? What if we were too abstract to task-level steps? For example, does a person read all the documents first then look at the questions; look at all the questions first then read the documents or something in between? These types of problem solving sequences are different from just looking at the ratios of the reading/scanning/scrolling strategies and could provide complementary data analysis that could provide valuable information to a student model.

One of the conscious decisions taken during each of these experiments was to try to provide a diverse set of documents that the students were required to read. People read different types of documents depending on the domain, their goals, even the stage of their lives [103] [104]. Perhaps their reading differs when reading different types of documents. Thus, the choices of reading material for our experiments included poetry, a technical article, response journals, songs, and newspaper articles. Do the students interact with them differently? Does the type of task change how they read these documents? Our results indicate that how a person is reading at a particular point within a document is an important predictor of a student's grade while working on a particular question about that document. Furthermore, we know from our experiments that a student can change their reading style based upon the level of cognitive difficulty of the problem they are currently working on and that this can affect the grade that they receive for answering the problem. However, we have not looked into to see if the student actually changed their reading behaviour based upon the type of document that they were currently reading or not.

In our first experiment, we set up a situation where memory and recall could possibly be examined to see if we could detect different behaviours when memory and recall situations arise. This involved a low level Bloom question from experiment 1A that was repeated in experiment 1B. However, there were more documents in 1B and the other questions in experiment 1B were different. For the one question that was the same, were there different behaviours from the same participant on experiment 1A compared to experiment 1B? A cursory examination of the results seem to indicate that there was a difference in detectable behaviours for some of the participants

that were involved in both experiment 1A and 1B. Although the data was collected and analyzed at a cursory level, a much more in-depth analysis would be required and might shed some interesting insights into memory and recall. Furthermore, a more explicit experiment would need to be designed, using more than one question, before any conclusions could be drawn.

Since our algorithm works as a single pass clustering algorithm, it does not need to consistently loop over and over the data taking up computational cycles. This means that the algorithm works efficiently and quickly and would be able to determine the ratios in near real-time. This poses an interesting question then about how early in the reading process can we detect that there is a problem with how a student is currently trying to solve a particular problem? In principle, it should be possible for an ALS to continuously check on a student's reading strategies, and to help a student during the reading and writing process rather than after the process has completed. This could involve modifying a student model based upon the student's reading style for a given level of difficulty. Given that past performance is often a strong indicator of future performance, knowing how well a student works at a given level of cognitive difficulty is something that both a student model and an ALS would find beneficial. Knowing where to direct the student so that they are given the opportunities to work on content areas that they are weak given their interaction past would help create a more personalized version of help for the student. Given the results demonstrated in this thesis, the types of help that we would be currently able to render would be limited to advice on how to read things differently. However, the long term possibilities, when combined with other discoveries and techniques hold promise. Determining exactly when there is enough data, and whether to intervene, however, are issues that will have to be explored through experimentation.

Students are not the only people involved in the learning process that could benefit from this research. For example, suppose a teacher was able to simply load some documents into a system, create some questions for the students based upon those documents and then highlight the portions of the document where the answers are located for the automated system. The system could then provide feedback to the student on how well they know the material within the set of documents without requiring direct teacher supervision and input. Such a system could even scaffold the level of difficulty of the questions to meet the student's style of reading.

Lastly, the most direct implication of this work is the possibility of automatically grading a student on a question based on how they solve that question. We know from this research that a student's reading style over a portion of the document is indicative of performance on a question. Will our predictions be accurate enough to altogether remove human judgment since we already know what a student's grade for answering questions at different levels of cognitive difficulty is going to be by how they read?

7.4 Long Range Implications

Reading comprehension itself is considered to be an ill-defined domain [105]. How a person reads a document, letter, poem, comic book, etc. and the context in which they read them do not fall into any well understood paradigm. Both in well-defined domains and ill-defined domains the requirement for reading is necessary for learning to occur. A tool that can analyze a student's reading behaviour and then provide feedback as to how successfully or unsuccessfully a student is learning as they move through the content would be invaluable, and would be a highly useful addition to the rather meagre array of tools currently available in ill-defined domains. Our approach holds out promise of providing a basis for the development of such a tool. As can be seen in section 7.2 and 7.3, there is still much research to be done, including of course all of the knowledge engineering issues of building an entire end application system for use by students. At the very least, however, our approach even as it stands could be used to inform student models for use by an ITS in an ill-defined domain or to provide direct feedback to students.

Additionally, the information from our system could be used to inform a teacher about problems a student may be having either with specific content or in their metacognitive skill set of reading comprehension skills. By metacognitive skill set we mean that the student is not choosing the appropriate reading style for the problem that they are currently trying to work on. This information alone can allow for a teacher to help the student correct a misconception or provide new skills for the student to learn.

The last area of education that our system could broadly impact is course management tools, such as Blackboard, Desire to Learn, WebCT, Moodle, Populi, etc., which have no intelligent help. The only corrections that a student gets from these types of systems are ones that

have been built specifically by the individual course designers and these are often trivial in nature in order to work within the constraints of the LCMS / LMS. Our work holds the promise of being able to eventually augment the capabilities of these systems by providing advice on how students read the material when solving problems, something that is a common activity across the diverse range of subjects supported by course management systems. Of course, this will require the teacher using the course management system to label the problems the students are to work on with their level of cognitive difficulty something that is often required of teachers even now without any automated tool.

By no means is this research definitive. Rather this is just the start of what will hopefully be something useful that will have a lasting impact. We have demonstrated that it is possible to analyze students' low level reading activities to find meaningful patterns that can inform the students, their teachers, or an intelligent system. We hope that this will not only prove useful, in and of itself, but also be suggestive of an approach that could be widely useful, especially in ill-defined domains.

References

[1] A. Head, M. Van Hoeck and D. Garson, "Lifelong learning in the digital age: A content analysis of recent research on participation," *First Monday,* vol. 20, no. 2, 2015.

[2] B. Stevens, "The story of the three Rs," *Education (14637073) Issue 307,* p. 6, 4 April 2008.

[3] S. Augustine and E. (Tranlator) Pusey, The Confessions of Saint Augustine, Gutenburg, 1901.

[4] R. Azevedo, D. Moos, A. Johnson and A. Chancey, "Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and Challenges," *Educational Psychologist,* vol. 45, no. 4, pp. 210-233, 2010.

[5] J. Mostow and W. Chen, "Generating Instruction Automatically for the Reading Strategy of Self-Questioning.," *Proceedings of the 14th International Conference on Artificial Intelligence in Education,* pp. 465-472, 2009.

[6] E. Thompson, A. Luxton-Reilly, J. Whalley, M. Hu and P. Robbins, "Bloom's Taxonomy for CS assesment.," *Proceedings of the Tenth Conference on Australasian Computing Education,* pp. 155-161, 2008.

[7] M. Zywno, "Hypermedia instruction and learning outcomes at different levels of Bloom's Taxonomy of Cognitive Domain," *Global Journal of Engineering Education,* vol. 7, no. 1, pp. 59-70, 2003.

[8] L. W. Anderson, D. R. Krathwohl and B. S. Bloom, Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, New York: Longman, 2000.

[9] B. Bloom, M. Englehart, M. Furst, E. Hill and D. Krathwohl, Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1 Cognitive Domain, New York: David McKay, 1956.

[10] R. Hilderman and T. Peckham, "A statistically sound alternative approach to mining contrast sets," *Proceedings of the 4th Australia Data Mining Conference,* pp. 157-172, 2005.

[11] C. Romero and S. Ventura, "Educational data mining: a review of the state of the art," *IEEE transactions on Systems, Man, and Cybernetics. Part C Applications and Reviews,* vol. 40, no. 6, pp. 601-618, 2010.

[12] G. Hammerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings.," *Proceedings of the ACM Conference on Information and Knowledge Management,* pp. 600-607, 2002.

[13] M. Alfassi, "Reading to learn: Effects of combined strategy instruction on high school students.," *Journal of Educational Research,* vol. 97, pp. 171-184, 2004.

[14] J. Beck, J. Mostow, A. Cuneo and J. Bey, "Can automated questioning help children's reading comprehension?," in *Proceedings of the tenth International Conference on Artificial Intelligence in Education,* Sidney, Australia, 2003.

[15] M. Sadoski and A. Paivio, "Toward a Unified Theory of Reading," *Scientific Studies of Reading,* vol. 11, pp. 337-365, 2007.

[16] A. Bandura, "Social Cognitive Theory: An Agentic Perspective," *Annual Review of Psychology,* vol. 52, no. 1, pp. 1-26, 2001.

[17] R. Ruddell and N. Unrau, "Reading as a meaning-construction process: The reader, the text, and the teacher," in *Theoretical Models and processes of reading,* Newark, International Reading Association, 1994, pp. 996-1056.

[18] D. McNamara, G. Jackson, C. Boonthum, K. Deng and X. He, "iSTART-ME: A natural language game-enhanced comprehension strategy tutor," *Journal of South China Normal University,* vol. 6, pp. 1-30, 2012.

[19] P. Carrell and J. Eisterhold, "Schema theory and ESL reading pedagogy," *TESOL Quarterly,* vol. 17, pp. 553-573, 1983.

[20] A. Anderson and T. Lynch, Listening, Oxford: Oxford University Press, 1998.

[21] M. Minsky, "A Framework for representing knowledge," *The psychology of computer vision,* pp. 211-277, 1975.

[22] L. Rosenblatt, "Towards a transactional theory of reading," *Journal of Reading Behavior,* vol. 1, no. 1, pp. 31-51, 1969.

[23] J. S. Brown, A. Collings and S. Duguid, "Situated cognition and the culture of learning," *Educational Researcher,* vol. 18, no. 1, pp. 32-42, 1989.

[24] A. Begg, "Enactivism and Some Implications for Education: A Personal Persepective," *Vinculum,* vol. 39, no. 2, pp. 4-12, 2002.

[25] Q. Li, B. Clark and I. Winchester, "Instructional Design and Technology Grounded in Enactivism: a Paradigm Shift?," *British Journal of Educational Technology,* vol. 41, no. 3, pp. 403-419, 2010.

[26] W. Clancey, Situated Cognition: On human knowledge and computer representations, New York: Cambridge University Press, 1997.

[27] J. Samuels, "Toward a theory of automatic information processing in reading, revisited," in *Theoretical models and processes of reading,* R. &. N. Unrau, Ed., Newark, DE, International Reading Association, 2004.

[28] J. Samuels, "Looking backwards: Reflections on a Career in Reading," *Journal of Literacy Research,* vol. 38, no. 3, pp. 327-344, 2006.

[29] T. Peckham and G. McCalla, "Mining Student Behavior Patterns in Reading Comprehension Task," *In Proceedings of the 5th International Conference on Educational Data Mining,* pp. 87-94, 2012.

[30] T. Peckham and G. McCalla, "Mining Reading Comprehension Within Educational Objective Frameworks," *Educational Data Mining,* 2014.

[31] W. Kintsch, "The role of knowledge in discourse comprehension: A construction-integration model," *Psychological Review,* vol. 95, pp. 163-182, 1988.

[32] W. Kintsch, "The construction-integration model of text comprehension and its implications for instruction," in *Theoretical models and processes of reading,* R. Ruddell and N. Unrau, Eds., Newark, DE., International Reading Association, 2004.

[33] M. Sadoski and A. Paivio, "A dual coding theoretical model of reading," in *Theoretical models and processes of reading,* R. R. &. N. Unrau, Ed., Newark, DE, International Reading Association, 2004.

[34] A. Paivio, "Dual coding theory: Retrospect and current status," *Candian Journal of Psychology,* vol. 45, no. 3, pp. 255-287, 1991.

[35] L. Uhr, "Teaching machine programs that generate problems as a function of interaction with students," in *Proceedings of the 24th National Conference ACM,* 1969.

[36] P. Suppes, "Some theoretical models for mathematics learning," *Journal of Research and Development in Education*, vol. 1, pp. 5-22, 1967.

[37] J. Carbonell, "AI in CAI: An artificial-intelligence approach to computer-assisted instruction," *IEEE Man-Machine Systems*, vol. 11, no. 4, pp. 190-202, 1970.

[38] D. Sleeman and J. Brown, "Introduction: Intelligent Tutoring Systems," in *Intelligent Tutoring Systems*, New York, Academic Press, 1982, pp. 1-11.

[39] M. Urban-Lurain, "Intelligent Tutoring Systems: An historic review in the context of the development of artifical intelligence and educational psychology," May 1996. [Online]. Available: http://www.cse.msu.edu/rgroups/cse101/ITS/its.htm.

[40] J. Greeno, A. Collings, B. Beranek and L. Resnick, "Cognition and learning," in *Handbook of educational psychology*, CEP 900 Course Pack, 1994, pp. 1-51.

[41] J. Self, "Student Models in computer-aided instruction," *International Journal of Man-machine studies*, vol. 6, no. 2, pp. 261-276, 1974.

[42] J. Brown and R. Burton, "Diagnostic models for procedural bugs in basic mathematical skills," *Cognitive Science*, vol. 2, pp. 155-192, 1978.

[43] R. Burton, "Diagnosing bugs in a simple procedural skill," in *Intelligent Tutoring Systems*, New York, Academic Press, 1982, pp. 157-183.

[44] J. Anderson, The Architecture of Cognition, Cambridge: Harvard University Press, 1983.

[45] K. Koedinger and J. Anderson, "Reifying Implicit Planning in Geometry: Guidelines for Model-Based Intelligent Tutoring Systems Design," in *Computers as Cognitive Tools*, Berlin, Springer-Verlag, 1993, pp. 15-45.

[46] A. Corbett and J. Anderson, "LISP Intelligent Tutoring System: Research in Skill Aquisition," in *Computer Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches*, Hillsdale, Lawrence Erlbaum Associates, 1992, pp. 73-109.

[47] R. Rosenberg, "A critical analysis of research on Intelligent Tutoring Systems," *Educational Technology*, vol. 27, no. 11, pp. 7-13, 1987.

[48] E. Wenger, Artificial Intelligence and Tutoring Systems: Computational and cognitive approaches to the communication of knowledge, San Francisco: Morgan Kaufmann, 1987.

[49] V. Shute and J. Psotka, "Intelligent Tutoring Systems: Past, Present and Future," in *Handbook of Research on Educational Communications and technology: Scholastic Publications*, Hillsdale, Lawrence Erlbaum Associates, 1996, pp. 570-600.

[50] M. Mark and J. E. Greer, "Evaluation Methodologies for Intelligent Tutoring Systems," *Journal of Artificial Intelligence in Education,* vol. 4, no. 2/3, pp. 129-153, 1999.

[51] W. Cooley and P. Lohnes, Evaluation Research in Education, New York: Irvington Publishers, 1976.

[52] S. D'Mello and A. Graesser, "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features," *User Modeling and User-Adapted Interaction,* vol. 20, no. 2, pp. 147-187, 2010.

[53] C. Epp and G. McCalla, "Protutor: Historic open learner modeling for pronounciation tutoring," *Artificial Intelligence in Education,* pp. 441-443, 2011.

[54] J. Mostow, J. Nelso-Taylor and J. Beck, "Comptuer-guided oral reading versus independent practice: Comparison of sustained silent reading to an automated reading tutor that listens," *Journal of Educational Computing Research,* vol. 49, no. 2, pp. 249-276, 2013.

[55] J. Mostow and G. Aist, "Giving help and praise in a reading tutor with imperfect listening," *CALICO Journal,* vol. 16, no. 3, pp. 407-424, 1999.

[56] A. Graesser, K. VanLehn, C. Rose, P. Jordan and D. Harter, "Intelligent Tutoring Systems with Conversational Dialogues," *AI Magazine,* vol. 22, no. 4, pp. 39-51, 2001.

[57] A. Graesser, N. K. Person and J. P. Magliano, "Collaborative Dialogue Patterns in Naturalistic One-on-One Tutoring," *Applied Cognitive Psychology,* vol. 9, no. 4, pp. 495-522, 1995.

[58] k. VanLen, P. Jordan, C. Rose, D. Behembe, M. Bottner, A. Gaydos and e. al., "The architecture of Why2-Atlas: A coach for qualitative physics essay writing," in *Intelligent tutoring Systems*, Heidelberg, Springer Berlin, 2002, pp. 158-167.

[59] J. Beck, J. Mostow, A. Cuneo and J. Bey, "Can automated questioning help children's reading comprehension?," *Proceedings of the 10th International Conference on Artificial Intelligence in Education,* pp. 380-382, 2003.

[60] C. Rose, "A Framework for Robust Semantic Interpretation," in *In Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linquistics*, San Francisco, Morgan Kaufmann, 2000, pp. 311-318.

[61] J. Brown and M. Eskenazi, "Retrieval of authentic documents for reader-specific lexical practice," in *InSTIL/ICALL Symposium*, 2004.

[62] M. Heilman and M. Eskenazi, "Self-assessment in vocabulary tutoring," *Intelligent Tutoring Systems*, pp. 656-658, 2008.

[63] K. Rosa and E. Maxine, "Self-Assessment in the REAP Tutor: Knowledge, Interest, Motivation & Learning," *International Journal of Artificial Intelligence in Education*, vol. 21, no. 4, pp. 237-253, 2013.

[64] D. McNamara, "SERT: Self-explanation reading training," *Discourse Processes*, vol. 38, pp. 1-30, 2004.

[65] T. O'Reilly, G. Sinclair and D. McNamara, "iSTART: A Web-based Reading Strategy Intervention That Improves Students' Science Comprehension," Lisbon, Portugal, 2004.

[66] J. Magliano, T. Trabasso and A. Graesser, "Strategic processing during comprehension," *Journal of Educational Psychology*, vol. 91, pp. 615-629, 1999.

[67] M. Chi, N. de Leeuw and L. C., "Eliciting self-explanations improves understanding," *Cognitive Science*, vol. 18, pp. 439-477, 1994.

[68] C. Bereiter and M. Bird, "Use of thinking aloud in identification and teaching of reading comprehension strategies," *Cognition and Instruction*, vol. 2, pp. 131-156, 1985.

[69] A. Palincsar and A. Brown, "Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities," *Cognition and Instruction*, vol. 1, pp. 117-175, 1984.

[70] A. Graesser and D. McNamara, "Scaffolding Deep Comprehension Strategies Through Point&Query, AutoTutor, and iSTART," *Educational Psychologist*, vol. 40, no. 4, pp. 225-234, 2005.

[71] W. Kintsch, Comprehension: A paradigm for cognition, New York: Cambridge University Press, 1998.

[72] D. McNamara, G. Jackson and A. Graesser, "Intelligent tutoring and games," *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, 2009.

[73] D. McNamara, T. Jackson, C. Boonthum, Y. Deng and X. He, "iSTART-ME: A Natural Language Game-enhanced Comprehension Strategy Tutor," *Journal of South China Normal University,* vol. 6, pp. 1-30, 2012.

[74] J. Mostow, S. Roth, A. Hauptmann and M. Kane, "A prototype reading coach that listens," *Proceedings of the 12th National Conference on Artficial Intelligence,* pp. 785-792, 1994.

[75] R. Allington, "If they don't read much, how they ever gonna get good?," *Journal of Reading,* vol. 21, pp. 57-61, 1977.

[76] G. Aist, "Helping Children Learn Vocabulary During Computer-Assisted Oral Reading," *Educational Technology and Society,* vol. 5, no. 2, 2002.

[77] G. Aist and J. Mostow, "Faster, better task choice in a reading tutor that listens," in *Speech Technology for Language Learning,* Netherlands, Swets & Zeitlinger Publishers, 2003.

[78] P. DooLittle, "Vygotsky's Zone of Proximal Development as a Theoretical Foundation for Cooperative Learning," *Journal on Excellence in College Teaching,* vol. 8, no. 1, pp. 83-103, 1997.

[79] A. C. Graesser and N. Person, "Question asking during tutoring," *American Education Research Journal,* vol. 31, pp. 104-137, 1994.

[80] T. Dillon, Questioning and teaching: A manual of practice, New York: Teachers College Press, 1988.

[81] A. Graesser, D. McNamara and K. VanLehn, "Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART," *Educational Psychologist,* vol. 40, no. 4, pp. 225-234, 2005.

[82] A. Mitrovic and A. Weerasinghe, "Revisiting ill-definedness and the consequences of ITSs.," *Artificial intelligence in education: Building learning systems that care from knowledge representation to affective modelling,* pp. 375-382, 2009.

[83] K. Koedinger, J. Anderson, W. Hadley and M. Mark, "Intelligent tutoring goes to school in the big city," *International journal of artificial intelligence in education,* vol. 8, pp. 30-43, 1997.

[84] A. Mitrovic, "Fifteen years of constraint-based tutors: What we have achieved and where we are going," *User modeling and user-adapted interaction,* vol. 22, no. 1, pp. 39-72, 2012.

[85] A. Mitrovic and S. Ohlsson, "Evaluation of a contraint-based tutor for database language," 1999.

[86] A. Graesser, K. VanLehn, C. Rose, P. Jordan and D. Harter, "Intelligent tutoring systems with conversational dialogues," *AI magazine,* vol. 22, no. 4, pp. 39-51, 2001.

[87] M. Mayo and A. Mitrovic, "Optimizing ITS behaviour with bayesian networks and decision theory," *Artificial Intelligence in Education,* vol. 12, no. 2, pp. 124-153, 2001.

[88] K. Rosa and M. Eskenazi, "Self-assessment in the REAP tutor: knowledge, interest, motivation & learning," *International Journal of Artificial Intelligence in Education,* vol. 21, no. 4, pp. 237-253, 2013.

[89] T. O'Reilly, G. Sinclair and D. McNamara, "iStart: A web-based reading strategy intervention that improves students science comprehension," *CELDA,* pp. 173-180, 2004.

[90] J. Alexander and A. Cockburn, "An empirical characterization of electronic document navigation," in *Proceedings of Graphics Interface 08*, Windsor, ON., Canada, 2008.

[91] K. Holmqvist, J. Holsanova, M. Barthelson and D. Lundqvist, "Reading or scanning? A study of newspaper and net paper reading," *Mind,* vol. 2, no. 3, pp. 657-670, 2003.

[92] J. Baker, " The impact of paging vs. scrolling on reading online text passages," *Usability News,* vol. 5, no. 1, pp. 2-5, 2003.

[93] M. Forehand, "Bloom's Taxonomy," in *Emerging perspectives on learning, teaching, and technology*, 2010, pp. 41-47.

[94] G. Schraw, "Measuring self-regulation in computer-based learning environments.," *Educational Psychologist,* vol. 45, no. 4, pp. 258-266, 2010.

[95] R. Hilderman and T. Peckham, "A statistically sound alternative approach to mining contrast sets," in *Proceedings of the 4th Australia Data Mining Conference*, Sydney, 2005.

[96] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory,* vol. 28, no. 2, pp. 129-137, 1982.

[97] T. Peckham and G. McCalla, "Mining student behavior patterns in reading comprehension tasks," in *Proceedings of the Fifth International Conference on Educational Data Mining,* Chania, Greece, 2012.

[98] E. Walker, A. Organ, V. Aleven and C. Jones, "Two approaches for providing adaptive support for discussion in an ill-defined domain," *Intelligent tutoring systems for ill-defined domains: Assessment and feedback,* 2008.

[99] D. Barton, M. Hamilton and R. Ivanic, Situated literacies: Reading and writing in context, Pyschology Press, 2000.

[100] A. Peacock, K. Xian and M. Wilkerson, "Typing patterns: A key to user identification," *IEEE Security & Privacy,* vol. 5, pp. 40-47, 2004.

[101] F. Wild, C. Stahl, G. Stermsek and G. Neumann, "Parameters driving effectiveness of automated essay scoring with LSA," in *Proceedings of 9th International Computer-Assisted Assessment Conference,* Loughborough, UK,, 2005.

[102] B. Morgan, W. Baggett and V. Rus, "Error Analysis as a Validation of Learning Progressions," *Educational Data Mining,* pp. 245-248, 2014.

[103] A. Jakobsen and K. Jensen, "Eye movement behaviour across four different types of reading task," *Copenhagen studies in language,* vol. 36, pp. 103-124, 2008.

[104] H. Distad, "A study of reading performance of pupils under different conditions on different types of materials," *Journal of Educational Psychology,* vol. 18, no. 4, p. 247, 1927.

[105] A. Rourke and J. Sweller, "The worked-example of using ill-defined problems: Learning to recognize designers' styles," *Learning and Instruction,* vol. 19, no. 2, pp. 185-199, 2009.

[106] B. Aworuwa and B. Nkoge, "The New Taxonomy of Educational Objectives and Implications for Designing Instruction for Distance Learning Delivery," *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education,* pp. 1394-1398, 2007.

[107] R. Marzano and J. Kendall, The new taxonomy of educational objectives, Corwin Press, 2007.

[108] B. Barron, D. Schwartz, N. Vye, A. Moore, A. Petronsino, L. Zech and J. Bransford, "Doing with understanding: Lessons from research on problem and project-based learning," *Journal of the Learning Sciences,* vol. 7, no. 3-4, pp. 271-311, 1998.

[109] R. Anderson and P. Pearson, "A schema-theoretic view of basic processes in reading comprehension," in *Handbook of reading research,* New York, Longman, 1984, pp. 829-864.

[110] S. Yao-Ting, C. Kuo-En and H. Jung-Sheng, "Improving children's reading comprehension and use of strategies through computer-based strategy training," *Computers in Human Behavior,* vol. 24, pp. 1552-1571, 2008.

[111] P. Wankat and F. Oreovicz, Teaching Engineering, New York: McGraw-Hill, 1993.

[112] J. Bransford, B. Stein, N. Nye, J. Franks, P. Auble, K. Merynski and G. Perfetto, "Differences in approaches to learning: An Overview.," *Journal of Experimental Psychology: General,* vol. 3, pp. 390-398, 1982.

[113] C. Fischer, "Revisiting the reader's rudder: A comprehension strategy.," *Journal of Adolescent & Adult Literacy,* vol. 47, pp. 248-256, 2003.

[114] S. Paris and A. Paris, "2001," *Classroom applications of research on self-regulated learning,* vol. 36, pp. 89-101, 2001.

[115] L. Razzaq, J. Patvarczki, S. Almeida, M. Vartak, F. Mingyu, N. Heffernan and K. Koedinger, "The ASSISTment Builder: Supporting the life cycle of tutoring systems content creation.," *Learning Technologies, IEEE Transactions on,* vol. 2, no. 2, pp. 157-166, 2009.

[116] H. Van Keer, "Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring," *British Journal of Educational Psychology,* vol. 74, pp. 37-70, 2004.

[117] R. Marzano, "Objectives that students understand.," *Educational Leadership,* vol. 68, no. 8, pp. 86-87, 2011.

[118] M. Wang, G. Haertel and H. Walberg, "Toward a knowledge base for school learning.," *Review of educational research,* vol. 63, no. 3, pp. 249-294, 1993.

[119] A. Yavuz and C. Cigdem, "Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011," *Computers & Education,* vol. 58, no. 2, pp. 835-842, 2012.

[120] K. Millis, J. Magliano and S. Todaro, "Measuring discourse-level processes with verbal protocols and latent semantic analysis," *Scientific Studies of Reading,* vol. 10, pp. 225-240, 2006.

[121] L. Rossenblatt, "The Transactional Theory of Reading and Writing," in *Theoretical models and processes of reading,* N. R.B. Ruddell, Ed., Newark, DE, International Reading Association, 2004.

[122] M. Sadoski and A. Paivio, Imagery and Text: A Dual Coding Theory of Reading and Writing, Mahwah, NJ: Erlbaum, 2001.

[123] K. VanLehn, A. Graesser, G. Jackson, A. Olney and C. Rose, "When is reading just as effective as one-on-one interactive human tutoring," in *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, 2005.

[124] V. Aleven, B. McLaren, J. Sewall and K. Koedinger, "The cognitive tutor authoring tools (ctat): Preliminary evaluation of efficiency gains," *Intelligent Tutoring Systems*, pp. 61-70, 2006.

[125] D. Berliner, "The science of psychology and practice of schooling: One hundred year journey of educational psychology from interst to distain, to respect for practice," in *American Psychological Association*, Washington, 1992.

[126] L. Pappano, *The year of the MOOC*, New York: New York Times, 2012.

[127] A. Meyer and J. Logan, "Taking the testing effect beyond the college feshman: Benefits for lifelong learning," *Psychology and aging*, vol. 28, no. 1, p. 142, 2013.

[128] K. VanLehn, A. Graesser and G. Jackson, "When is reading just as effective as one-on-one interactive human tutoring," 2005.

[129] J. Mostow, "Generating instruction automatically for the reading strategy of self-questioning," 2009.

[130] V. Aleven, B. McLaren, J. Sewall and K. Koedinger, "The cognitive tutor authoring tools (ctat): Preliminary evaluation of efficiency gains," *Intelligent Tutoring Systems*, pp. 61-70, 2006.

[131] C. Conati, "Intelligent tutoring systems: New challenges and directions," *IJCAI*, vol. 9, pp. 2-7, 2009.

[132] K. Collins-Tompson and J. Callan, "Information Retrival for Language Tutoring: An Overview of the REAP Project," *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 544-545, 2004.

[133] V. Aleven, K. Koedinger and K. Cross, "Tutoring answer explanations fosters learning with understanding," *Artificial intelligence in education*, pp. 199-206, 1999.

[134] G. Weber and P. Brusilovsky, "ELM-ART: an adaptive versatile system for web-based instruction," *International journal of artificial intelligence in education,* vol. 12, pp. 351-384, 2001.

[135] I. Goldin, R. Ashley and R. Pinkus, "Teaching case analysis thorugh framing: prospectsfor an ITS in an ill-defined domain," *Workshop on AIED applications in ill-defined domains,* pp. 83-91, 2008.

[136] D. Kuhn, "Cognitive Development," in *Developmental Psychology: An Advanced Textbook*, Hillsdale, Lawrence Erlbaum, 1992, pp. 211-272.

[137] T. Clarke, "A natural history of the mind: A guide for cognitive science.," *Behavioral and Brain Science,* vol. 16, no. 4, pp. 754-755, 1993.

[138] W. Sack, E. Soloway and P. Weingrad, "Re-writing cartesian student models," in *Student modeling: The key to individualized knowledge-based instruction*, Berlin, Springer, 1994, pp. 355-376.

[139] A. Mitrovic, "An intelligent SQL tutor on the web," *International Journal of Artificial Intelligence in Education,* vol. 13, no. 2, pp. 173-197, 2003.

[140] A. Peacock, X. Ke and M. Wilkerson, "Typing Patterns: A key to user identification," *IEEE Security & Privacy,* vol. 2, no. 5, pp. 40-47, 2004.

[141] G. McCalla, "The ecological approach to the design of e-learning environments: Purpose-based capture and use of information about learners," *Journal of Interactive Media,* vol. 2004, no. 1, pp. Art-3, 2010.

Appendix 1 Experiment 1

Experiment 1A

Question 1 How many of the 11 recommendations made by the Canadian Privacy Commission were accepted by FaceBook?

Answer 1: Section 1 Collection of Date of Birth (2), Section 2 Default Privacy Settings (2), Section 3 Facebook Advertising(1), Section 7a Account Deactivation and Deletion (1), Section 10 Monitoring for Anomalous Activity(1). There were 7 different recommendations that were accepted by facebook from 5 different categories

Question 2 Identify the two main findings with Facebook allowing third-party applications to access private data?

Answer2: 1)Facebook had inadequate safeguards to effectively restrict these outside developers from accessing users' profile information, along with information about their online friends. 2) Facebook was not obtaining users' meaningful consent to the disclosure of their personal information to application developers when either they or their friends add applications.

Question 3 Discuss if Facebook collects information from other sources than Facebook?

Answer3: Facebook's privacy policy does contain language about collecting personal information from outside sources, in fact, it does not do so at the present time.

Question 4 Was Facebook involved in misrepresenting itself by claiming to be a purely social networking site when in fact it was engaged in other activities?

Answer4: Section 11 Deception and Misrepresentation. There was no evidence of intent to deceive or misrepresent.

Experiment 1B

Question 1 How many of the 11 recommendations made by the Canadian Privacy Commission against Facebook were rejected?

Answer 1: Section 1 Collection of Date of Birth (2), Section 2 Default Privacy Settings (2), Section 3 Facebook Advertising(1), Section 7a Account Deactivation and Deletion (1), Section 10 Monitoring for Anomalous Activity(1). There were 7 different recommendations that were accepted by facebook from 5 different categories

Question 2 Identify the two main findings with Facebook allowing third-party applications to access private data?

Answer2: 1)Facebook had inadequate safeguards to effectively restrict these outside developers from accessing users' profile information, along with information about their online friends. 2) Facebook was not obtaining users' meaningful consent to the disclosure of their personal information to application developers when either they or their friends add applications.

Question 3 Discuss if Facebook collects information from other sources other than Facebook?

Answer3: Facebook's privacy policy does contain language about collecting personal information from outside sources, in fact, it does not do so at the present time.

Question 4 Critique Facebook's use of third party applications.

Answer4: Multiple answers… use rubric to mark

Question 5 Explain how to stop someone from posting on your wall.

Answer5: Visit the Profile Privacy settings page. There is a section labelled Wall Posts. From here you can completely disable your friend's ability to post on your wall. You can also select a specific friend list that can post on your wall.

Question 6 Express, using both personal information and information from the articles, whether or not Facebook has done enough to ensure your privacy.

Answer6: Multiple answers… use rubric to mark

Question 7 Discuss the policy about Facebook retaining the right to keep any payment or transaction details performed on Facebook.

Answer7: Facebook by default will store details of transactions made from the Facebook site. You may remove this transaction information by visiting the payments page. The process of storing these details is that further transactions can be made more expediently if the information is retained from previous transactions. However, there must be sufficient safeguards in place to make sure that this information is not accessed by unauthorized parties. Recent exploits of data being accessed from Sony is a current caution that should be noted prior to letting any corporation store your information.

Question 8 Choose a side and debate if the Age policy for Facebook usage is fair.

Answer8: Multiple answers… use rubric to mark

Statistics

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	1.084	3	0.361	79.94	3.14E-16	88.55
within groups	0.167	37	0.004521			11.45
total	1.252	40				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.44279	0.604833	0.739092	0.88751	-	-
Gabriel comparison interval	0.042	0.044	0.038	0.042	-	-
n	10	9	12	10	-	-

Statistics for Bloom Level 1

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.08311	0.07745	0.08089
60,30,10	0.16204*	-	0.07976	0.08311
70,20,10	0.2963*	0.13426*	-	0.07745
80,10,10	0.4447*	0.28268*	0.14842*	-

Tukey-Kramer Bloom Level

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	2.993	3	0.998	39.308	3.74E-11	80.56
within groups	0.863	34	0.025			19.44
total	3.857	37				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.239538	0.45295	0.73015	0.911982	-	-
Gabriel comparison interval	0.087	0.128	0.111	0.095	-	-
n	13	6	8	11	-	-

Statistics for Bloom Level 2

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.21238	0.19337	0.17629
60,30,10	0.21341*	-	0.2324	0.21839
70,20,10	0.4906*	0.2772*	-	0.19995
80,10,10	0.6724*	0.459*	0.18183	-

Tukey-Kramer for Bloom Level 2

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	0.82	2	0.41	147.929	4.80E-11	95.91
within groups	0.044	16	0.002773			4.09
total	0.865	18				
	50,30,20	**70,20,10**	**80,10,10**			
mean	0.45946	0.796286	0.988714	-	-	-
Gabriel comparison interval	0.044	0.037	0.037	-	-	-
n	5	7	7	-	-	-

Statistics for Bloom Level 3

	50,30,20	70,20,10	80,10,10
50,30,20	-	0.07956	0.07956
70,20,10	0.3368*	-	0.07263
80,10,10	0.5293*	0.19243*	-

Tukey-Kramer for Bloom Level 3

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	1.134	3	0.378	25.559	0.000029	88.33
within groups	0.163	11	0.015			11.67
total	1.297	14				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.00E00	0.67886	0.46995	0.985871	-	-
Gabriel comparison interval	0.27	0.121	0.191	0.102	-	-
n	1	5	2	7	-	-

Statistics for Bloom Level 5

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.40104	0.44838	0.39138
60,30,10	0.6789*	-	0.3063	0.21437
70,20,10	0.47*	0.20891	-	0.29353
80,10,10	0.9859*	0.30701*	0.5159*	-

Tukey-Kramer for Bloom Level 5

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	0.165	1	0.165	50.766	0.000385	94.31
within groups	0.02	6	0.003254			5.69
total	0.185	7				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	-	0.63	-	0.961833	-	-
Gabriel comparison interval	-	0.00E0	-	0.00E0	-	-
n	-	2	-	6	-	-

Statistics for Bloom Level 6

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20				
60,30,10		-		0
70,20,10				
80,10,10		0.3318*		-

Tukey-Kramer for Bloom Level 6

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	115.637	3	38.546	1.401	0.246	-
within groups	3438.809	125	27.51			-
total	3554.446	128				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.339283	3.080867	0.731869	0.941936	-	-
Gabriel comparison interval	1.84	2.022	1.84	1.445	-	-
n	29	24	29	47	-	-

Statistics for Clusters with No Bloom Analysis

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	3.76929	3.58709	3.22541
60,30,10	2.7416	-	3.76929	3.4269
70,20,10	0.3926	2.349	-	3.22541
80,10,10	0.6027	2.1389	0.21007	-

Tukey-Kramer for Clusters with No Bloom Analysis

Documents

Canadian Allegations

***Section 1 – Collection of Date of Birth

1. That Facebook was unnecessarily requiring users to provide their dates of birth as a condition of registration, in contravention of Principle 4.3.3.

2. That Facebook was not adequately explaining to users why they had to provide their dates of birth and how these would be used, in contravention of Principle 4.3.2.

>>> Findings:

1. Date of birth is acceptable as a condition of service since purposes for its use are appropriate.

2. However, Facebook was not clearly explaining these purposes.

^^^ Recommendation(s):

· Facebook was asked to clearly tell users, when registering, why birth dates are required.

· It was also asked to clarify in its site documentation the reasons for collecting date of birth and how it may be used.

/// Response:

Facebook agreed to all recommendations.

Conclusion: Well-founded and resolved

***Section 2 – Default Privacy Settings

1. That Facebook, by preselecting default privacy settings, was in effect using opt-out consent for the use and disclosure of personal information without meeting the requirements for opt-out consent as articulated in previous findings of our Office. Specifically, it was contended that much of the personal information being shared by users, including photographs, marital status, age, and hobbies, is sensitive and therefore requires express consent.

2. That Facebook does not, in the context of its privacy settings, make a reasonable effort to advise users of the purposes for which and the extent to which their personal information is used and disclosed. Specifically,

· Facebook does not inform users of the extent to which their personal information may be shared through the default settings and so does not have meaningful consent.

· Facebook does not direct users to the privacy settings when they complete registration, when they upload photos, or when Facebook makes changes to the settings.

· Facebook does not inform users that failure to alter the default settings constitutes consent to those settings.

· Facebook fails to provide adequate notice to users posting photo albums that the default privacy settings for photo albums enable sharing with everyone, with the result that a user's non-friends can view his or her photographs and associated comments, even if the user's profile is searchable only by his or her friends.

· When users sign up for a network, their default privacy settings enable the sharing of their personal information, including sensitive information, with everyone on the network.

>>> Findings:

1. Users voluntarily upload their personal information for the purpose of sharing it with others.

2. Default privacy settings are acceptable as long as they meet users' reasonable expectations. They do not in two instances: photo albums (set to "Everyone") and search (consent to being searchable by search engines).

3. Sufficient information was not provided to users with regard to how privacy settings are defaulted and the implications of not modifying the defaulted settings.

^^^ Recommendations:

Facebook was asked to:

· make user profiles inaccessible to search engines by default;

· change the default setting for photo albums to "Your Networks and Friends," and

· provide a link to the privacy settings at registration, accompanied by a statement of what the settings are for, that Facebook has preselected settings, and that settings can be changed according to preferences.

/// Response:

Facebook is making changes to its privacy settings a) by allowing users to choose a high, medium, low setting and b) introducing a per-object privacy that allows users to choose privacy settings on individual photos and pieces of content such as status updates.

Conclusion: Well-founded and resolved

***Section 3 – Facebook Advertising

1. That Facebook was not making a reasonable effort to notify users clearly that it used their personal information for advertising purposes, in violation of Principle 4.3.2.

2. That Facebook, for Social Ads in particular, was improperly using opt-out rather than opt-in consent in accordance with Principle 4.3.6, given the sensitivity of users' personal information.

3. That Facebook was not allowing users to opt out of Facebook Ads, in contravention of Principle 4.3.8.

4. Since users were not allowed to opt out of Facebook Ads, Facebook was unnecessarily requiring users to agree to such ads as a condition of service, in violation of Principle 4.3.3.

>>> Findings:

1. Users cannot opt out of all advertising as advertising revenues are required to run site (which is free to users).

2. Users can opt out of Social Ads - this type of advertising is more intrusive (the individual is used to promote products, services, etc.) and therefore users should not be required to consent to Social Ads.

3. Requiring users to consent to Facebook Ads is acceptable as they are not being co-opted into endorsing a product.

4. However, Facebook is not informing users of advertising purposes.

^^^ Recommendations:

· Facebook was asked to expand the advertising section of the Privacy Policy to more fully explain advertising and to inform users that their profile information is used for targeted advertising.

/// Response:

Facebook agreed to describe advertising more clearly and to configure its systems to allow users to more easily find information about advertising.

Conclusion: Well-founded and resolved

***Section 4 – Third-Party Applications

1. That Facebook was not informing users of the purpose for disclosing personal information to third-party application developers, in contravention of Principles 4.2.2 and 4.2.5.

2. That Facebook was providing third-party application developers with access to personal information beyond what was necessary for the purposes of the application, in contravention of Principle 4.4.1.

3. That Facebook was requiring users to consent to the disclosure of personal information beyond what was necessary to run an application, in contravention of Principle 4.3.3.

4. That Facebook was not notifying users of the implications of withdrawing consent to sharing personal information with third-party application developers, in contravention of Principle 4.3.8.

5. That Facebook was allowing third-party application developers to retain a user's personal information after the user deleted the application, in contravention of Principle 4.5.3.

6. That Facebook was allowing third-party developers access to the personal information of users when their friends or fellow network members added applications without adequate notice, in contravention of Principle 4.3.2.

7. That Facebook was not adequately safeguarding personal information in that it was not monitoring the quality or legitimacy of third-party applications or taking adequate steps against inherent vulnerabilities in many programs on the Facebook Platform, in contravention of Principle 4.7.

8. That Facebook was not effectively notifying users of the extent of personal information that is disclosed to third-party application developers and was providing users with misleading and unclear information about sharing with third-party application developers, in contravention of Principles 4.3.and 4.8.

9. That Facebook was not taking responsibility for the personal information transferred to third-party developers for processing, in contravention of Principle 4.1.3.

10. That Facebook was not permitting users to opt out of sharing their name, networks, and friend lists when their friends add applications, in contravention of Principle 4.3 and subsection 5(3).

>>> Findings:

1. Facebook had inadequate safeguards to effectively restrict these outside developers from accessing users' profile information, along with information about their online friends.

2. Facebook was not obtaining users' meaningful consent to the disclosure of their personal information to application developers when either they or their friends add applications.

^^^ Recommendations:

· Facebook was asked to implement technological measures to limit application developers' access to user information that is not required to run a specific application.

· The site should also ensure that users are informed of the specific information that an application requires and for what purpose. In addition, users' express consent for the developer's access to the specific information must be sought each time someone signs up for an application.

· Finally, measures are needed to prohibit all disclosure of the personal information of users who are not themselves adding an application.

/// Response:

Facebook has not agreed to implement the recommendations.

Conclusion: Well-founded

***Section 5 – New Uses of Personal Information

1. That Facebook was not notifying users of new purposes for which their personal information would be collected, used, or disclosed, in violation of Principle 4.2.4.

>>> Findings:

There was no evidence that Facebook had failed to inform its users of new uses.

/// Conclusion: Not well-founded

***Section 6 – Collection of Personal Information from Sources Other than Facebook

1. That Facebook was failing to provide users with specific information relating to the purposes and method of collecting personal information from sources outside Facebook, the sources of the information, and the use and disclosure of the information.

2. Having failed to inform users of these specifics, Facebook was therefore not obtaining their meaningful consent.

>>> Findings:

Although Facebook's privacy policy contains language about collecting personal information from outside sources, in fact, it does not do so at the present time.

/// Conclusion: Not well-founded

***Section 7(a) – Account Deactivation and Deletion

1. That Facebook was offering only an account deactivation option as distinct from an account deletion option and was therefore inappropriately depriving users of a means whereby they could delete all their personal information from the site.

>>> Findings:

1. Account deactivation and deletion are explained on the site, but not in the same part of the site. It may cause some users to believe that deactivation is their only option.

2. It is retaining personal information from deactivated accounts indefinitely.

^^^ Recommendations:

· Facebook was asked to develop, institute and inform users about a retention policy under which the personal information of users who have deactivated their accounts will be deleted from Facebook's servers after a reasonable length of time.

· As a best practice, the Assistant Commissioner also suggested that Facebook make the account deletion option more prominent for users.

/// Response:

Facebook agreed to add information about account deletion to its privacy policy, but declined to develop a retention policy for deactivated accounts.

Conclusion: Well-founded

***Section 7(b) – Accounts of Deceased Users

1. By including only in its Terms of Use and not in its Privacy Policy a notice of its intention to keep deceased users' profiles active for memorial purposes, Facebook was not obtaining users' meaningful consent for such use of their personal information.

2. That Facebook was obligating users, in contravention of Principle 4.3.3, to consent to this purpose as a condition of service even though memorializing a profile is not necessary to Facebook's primary purpose of providing a social networking venue.

>>> Findings:

1. Memorialization can be considered a primary purpose since most users would reasonably expect it.

2. However, users are not informed of the practice, whereby they would effectively provide their consent to it.

^^^ Recommendation:

· Facebook was asked to include in its Privacy Policy an explanation about the practice of using the personal information to memorialize the accounts of deceased users.

/// Response:

Facebook did not agree to implement the recommendation, considering it unnecessary under the law.

Conclusion: Well-founded

***Section 8 – Personal Information of Non-users

1. That Facebook was not obtaining consent from non-users for the uploading of their personal information to the site, in contravention of Principle 4.3, in the following situations:

· Users can post the personal information of non-users in their own profiles, as well as the profiles of other users through features such as "News Feed" and "Wall". Also, users can tag images of non-users with their names in photos or videos.

· Users can provide Facebook with the email addresses of non-users for the purpose of inviting them to join the site.

>>> Findings:

1. When users post personal information about non-users on walls, profiles, or the News Feed, such postings are made for personal purposes only and fall outside the scope of the Act.

2. In the cases of tagging of and invitations to non-users, the Act only applies where Facebook uses non-users personal information for purposes of its own, namely, informing non-users when they have been tagged or inviting them to join Facebook.

3. Facebook may rely on users to obtain the consent of non-users for these two purposes, provided that the company exercises reasonable due diligence. This could simply mean taking steps to ensure that users know that they must obtain non-users' consent before disclosing their email addresses to Facebook, and punishing users who violate the consent requirement.

4. However, such information is currently missing from the Privacy Policy.

^^^ Recommendations:

· It was asked to implement measures to improve the invitation feature so as to address our concerns about non-users' lack of knowledge and consent to Facebook's collection, use, and retention of their email addresses; and

· It was asked to set a reasonable time limit on the retention of non-users' email addresses after they have been invited to join Facebook.

/// Response:

Facebook declined to implement the first and second recommendations above on grounds that the site already provides "significantly greater notice to non-users as to the presence of any information about them on our site than does any other site on the web."

Facebook also noted that it could not realistically delete the personal information of non-users when it is uploaded by users, because that information is in the user's possession and control. As such, non-user data is the responsibility of the user who uploads it.

Facebook made no direct response to the third recommendation.

Conclusion: Well-founded

***Section 9 – Facebook Mobile and Safeguards

1. With respect to users of the mobile version of the Facebook website (Mobile Facebook), it was alleged that, by providing such users with a persistent cookie having no apparent expiration date, Facebook was failing to properly safeguard their personal information, in contravention of Principles 4.7, 4.7.1, and 4.7.3.

2. Specifically, CIPPIC cited the following security concerns:

1. If a user logs onto his or her Facebook account by means of another person's mobile device and forgets to log off, the other person will have access to the user's Facebook account indefinitely, even if the user changes the password.

2. If a user gives his or her Facebook password to another person, that person can log in as the user on a mobile device and have access indefinitely, even if the user changes the password.

3. In CIPPIC's view, Facebook should have a cookie that expires within an appropriate period of time and whenever users change their passwords online.

>>> Findings:

1. Facebook uses a persistent cookie with a 14-day expiration date. A password change on another platform causes a session open on Facebook Mobile to close and require re-authentication for a user to log back on.

2. Therefore, Facebook provides users with a simple method of logging out of sessions on Facebook Mobile, as well as the ability to effectively cease Facebook sessions initiated on mobile devices by changing their passwords on other platforms.

3. As such, Facebook provides users of Facebook Mobile with adequate safeguards to protect their sessions from unauthorized access.

/// Conclusion: Not well-founded

***Section 10 – Monitoring for Anomalous Activity

1. That Facebook was not informing users that it monitors the site for anomalous behaviour and, in particular, failed to mention this practice in its Privacy Policy, in violation of Principle 4.8.

>>> Findings:

1. The practice of monitoring the site for anomalous behaviour was appropriate, but Facebook was not making a reasonable effort to inform users of it.

^^^ Recommendation:

· Facebook was asked to explain the practice in its Privacy Policy

///Response:

Facebook agreed to the recommendation.

Conclusion: Well-founded and resolved

***Section 11 – Deception and Misrepresentation

1. That Facebook was misrepresenting itself by claiming to be purely a social networking site when in fact it was engaged in other activities not clearly explained, such as advertising and third-party applications, in contravention of Principles 4.3.2 and 4.4.2.

2. That Facebook was misrepresenting users' level of control over their personal information, in contravention of Principles 4.3.2 and 4.4.2.

>>> Findings:

There was no evidence of intent to deceive or misrepresent.

Conclusion: Not well-founded

Facebook Privacy Policy

Date of last revision: December 22, 2010.

This policy contains nine sections, and you can jump to each by selecting the links below:

1. Introduction

2. Information We Receive

3. Sharing information on Facebook

4. Information You Share With Third Parties

5. How We Use Your Information

6. How We Share Information

7. How You Can Change or Remove Information

8. How We Protect Information

9. Other Terms

1. Introduction

Safe Harbor. Facebook also complies with the EU Safe Harbor framework as set forth by the Department of Commerce regarding the collection, use, and retention of data from the European Union. As part of our participation in the Safe Harbor, we agree to resolve all disputes you have with us in connection with our policies and practices through TRUSTe. We will also provide initial responses to access requests within a reasonable period of time. To view our certification, visit the U.S. Department of Commerce's Safe Harbor Web site.

Scope. This privacy policy covers all of Facebook. It does not, however, apply to entities that Facebook does not own or control, such as applications and websites using Platform. By using or accessing Facebook, you agree to our privacy practices outlined here.

No information from children under age 13. If you are under age 13, please do not attempt to register for Facebook or provide any personal information about yourself to us. If we learn that we have collected personal information from a child under age 13, we will delete that information as quickly as possible. If you believe that we might have any information from a child under age 13, please contact us through this help page.

Parental participation. We strongly recommend that minors 13 years of age or older ask their parents for permission before sending any information about themselves to anyone over the Internet and we encourage parents to teach their children about safe internet use practices. Materials to help parents talk to their children about safe internet use can be found on this help page.

2. Information We Receive

Information you provide to us:

Information About Yourself. When you sign up for Facebook you provide us with your name, email, gender, and birth date. During the registration process we give you the opportunity to connect with your friends, schools, and employers. You will also be able to add a picture of yourself. In some cases we may ask for additional information for security reasons or to provide specific services to you. Once you register you can provide other information about yourself by connecting with, for example, your current city, hometown, family, relationships, networks, activities, interests, and places. You can also provide personal information about yourself, such as your political and religious views.

Content. One of the primary reasons people use Facebook is to share content with others. Examples include when you update your status, upload or take a photo, upload or record a video, share a link, create an event or a group, make a comment, write something on someone’s Wall, write a note, or send someone a message. If you do not want us to store metadata associated with

content you share on Facebook (such as photos), please remove the metadata before uploading the content.

Transactional Information. We may retain the details of transactions or payments you make on Facebook. If you do not want us to store your payment source account number, you can remove it using your payments page.

Friend Information. We offer contact importer tools to help you upload your friends' addresses so that you can find your friends on Facebook, and invite your contacts who do not have Facebook accounts to join. If you do not want us to store this information, visit this help page. If you give us your password to retrieve those contacts, we will not store your password after you have uploaded your contacts' information.

Information we collect when you interact with Facebook:

Site activity information. We keep track of some of the actions you take on Facebook, such as adding connections (including joining a group or adding a friend), creating a photo album, sending a gift, poking another user, indicating you "like" a post, attending an event, or connecting with an application. In some cases you are also taking an action when you provide information or content to us. For example, if you share a video, in addition to storing the actual content you uploaded, we might log the fact that you shared it.

Access Device and Browser Information. When you access Facebook from a computer, mobile phone, or other device, we may collect information from that device about your browser type, location, and IP address, as well as the pages you visit.

Cookie Information. We use "cookies" (small pieces of data we store for an extended period of time on your computer, mobile phone, or other device) to make Facebook easier to use, to make our advertising better, and to protect both you and Facebook. For example, we use them to store your login ID (but never your password) to make it easier for you to login whenever you come back to Facebook. We also use them to confirm that you are logged into Facebook, and to know when you are interacting with Facebook Platform applications and websites, our widgets and Share buttons, and our advertisements. You can remove or block cookies using the settings in your browser, but in some cases that may impact your ability to use Facebook.

Information we receive from third parties:

Facebook Platform. We do not own or operate the applications or websites that you use through Facebook Platform (such as games and utilities). Whenever you connect with a Platform application or website, we will receive information from them, including information about actions you take. In some cases, in order to personalize the process of connecting, we may receive a limited amount of information even before you connect with the application or website.

Information from other websites. We may institute programs with advertising partners and other websites in which they share information with us:

We may ask advertisers to tell us how our users responded to the ads we showed them (and for comparison purposes, how other users who didn't see the ads acted on their site). This data sharing, commonly known as "conversion tracking," helps us measure our advertising effectiveness and improve the quality of the advertisements you see.

We may receive information about whether or not you've seen or interacted with certain ads on other sites in order to measure the effectiveness of those ads.

If in any of these cases we receive data that we do not already have, we will "anonymize" it within 180 days, meaning we will stop associating the information with any particular user. If we institute these programs, we will only use the information in the ways we explain in the "How We Use Your Information" section below.

Information from other users. We may collect information about you from other Facebook users, such as when a friend tags you in a photo, video, or place, provides friend details, or indicates a relationship with you.

3. Sharing information on Facebook.

This section explains how your privacy settings work, and how your information is shared on Facebook. You should always consider your privacy settings before sharing information on Facebook.

Name and Profile Picture. Facebook is designed to make it easy for you to find and connect with others. For this reason, your name and profile picture do not have privacy settings. If you are uncomfortable with sharing your profile picture, you should delete it (or not add one). You can also control who can find you when searching on Facebook or on public search engines using the Applications and Websites privacy setting.

Contact Information. Your contact information settings control (available when customizing your privacy settings) who can contact you on Facebook, and who can see your contact information such as your email and phone number(s). Remember that none of this information is required except for your email address, and you do not have to share your email address with anyone.

Personal Information. Your personal information settings control who can see your personal information, such as your religious and political views, if you choose to add them. We recommend that you share this information using the friends of friends setting.

Posts by Me. You can select a privacy setting for every post you make using the publisher on our site. Whether you are uploading a photo or posting a status update, you can control exactly who can see it at the time you create it. Whenever you share something look for the lock icon. Clicking on the lock will bring up a menu that lets you choose who will be able to see your post. If you decide not to select your setting at the time you post the content, your content will be shared consistent with your Posts by Me default privacy

(available when customizing your privacy settings).

Gender and Birth Date. In addition to name and email address, we require you to provide your gender and birth date during the registration process. We ask for your date of birth to verify that you are 13 or older, and so that we can better limit your access to content and advertisements that are not age appropriate. Because your date of birth and gender are required, you cannot delete them. You can, however, edit your profile to hide all (or part) of such fields from other users.

Other. Here are some other things to remember:

Some of the content you share and the actions you take will show up on your friends' home pages and other pages they visit.

If another user tags you in a photo or video or at a place, you can remove the tag. You can also limit who can see that you have been tagged on your profile from your privacy settings.

Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy settings, or it was copied or stored by other users.

You understand that information might be reshared or copied by other users.

Certain types of communications that you send to other users cannot be removed, such as messages.

When you post information on another user's profile or comment on another user's post, that information will be subject to the other user's privacy settings.

If you use an external source to publish information to Facebook (such as a mobile application or a Connect site), you should check the privacy setting for that post, as it is set by that external source.

"Everyone" Information. Information set to "everyone" is publicly available information, just like your name, profile picture, and connections. Such information may, for example, be accessed by everyone on the Internet (including people not logged into Facebook), be indexed by third party search engines, and be imported, exported, distributed, and redistributed by us and others without privacy limitations. Such information may also be associated with you, including your name and profile picture, even outside of Facebook, such as on public search engines and

when you visit other sites on the internet. The default privacy setting for certain types of information you post on Facebook is set to "everyone." You can review and change the default settings in your privacy settings. If you delete "everyone" content that you posted on Facebook, we will remove it from your Facebook profile, but have no control over its use outside of Facebook.

Minors. We reserve the right to add special protections for minors (such as to provide them with an age-appropriate experience) and place restrictions on the ability of adults to share and connect with minors, recognizing this may provide minors a more limited experience on Facebook

4. Information You Share With Third Parties.

Facebook Platform. As mentioned above, we do not own or operate the applications or websites that use Facebook Platform. That means that when you use those applications and websites you are making your Facebook information available to someone other than Facebook. Prior to allowing them to access any information about you, we require them to agree to terms that limit their use of your information (which you can read about in Section 9 of our Statement of Rights and Responsibilities) and we use technical measures to ensure that they only obtain authorized information. To learn more about Platform, visit our About Platform page.

Connecting with an Application or Website. When you connect with an application or website it will have access to General Information about you. The term General Information includes your and your friends' names, profile pictures, gender, user IDs, connections, and any content shared using the Everyone privacy setting. We may also make information about the location of your computer or access device and your age available to applications and websites in order to help them implement appropriate security measures and control the distribution of age-appropriate content. If the application or website wants to access any other data, it will have to ask for your permission.

We give you tools to control how your information is shared with applications and websites that use Platform. For example, you can block all platform applications and websites completely or block specific applications from accessing your information by visiting your Applications and Websites privacy setting or the specific application's "About" page. You can also use your privacy settings to limit which of your information is available to "everyone".

You should always review the policies of third party applications and websites to make sure you are comfortable with the ways in which they use information you share with them. We do not guarantee that they will follow our rules. If you find an application or website that violates our rules, you should report the violation to us on this help page and we will take action as necessary.

When your friends use Platform. If your friend connects with an application or website, it will be able to access your name, profile picture, gender, user ID, and information you have shared with "everyone." It will also be able to access your connections, except it will not be able to

access your friend list. If you have already connected with (or have a separate account with) that website or application, it may also be able to connect you with your friend on that application or website. If the application or website wants to access any of your other content or information (including your friend list), it will have to obtain specific permission from your friend. If your friend grants specific permission to the application or website, it will generally only be able to access content and information about you that your friend can access. In addition, it will only be allowed to use that content and information in connection with that friend. For example, if a friend gives an application access to a photo you only shared with your friends, that application could allow your friend to view or print the photo, but it cannot show that photo to anyone else.

We provide you with a number of tools to control how your information is shared when your friend connects with an application or website. For example, you can use your Application and Websites privacy setting to limit some of the information your friends can make available to applications and websites. You can block all platform applications and websites completely or block particular applications or websites from accessing your information. You can use your privacy settings to limit which friends can access your information, or limit which of your information is available to “everyone.” You can also disconnect from a friend if you are uncomfortable with how they are using your information.

Pre-Approved Third-Party Websites and Applications. In order to provide you with useful social experiences off of Facebook, we occasionally need to provide General Information about you to pre-approved third party websites and applications that use Platform at the time you visit them (if you are still logged in to Facebook). Similarly, when one of your friends visits a pre-approved website or application, it will receive General Information about you so you and your friend can be connected on that website as well (if you also have an account with that website). In these cases we require these websites and applications to go through an approval process, and to enter into separate agreements designed to protect your privacy. For example, these agreements include provisions relating to the access and deletion of your General Information, along with your ability to opt-out of the experience being offered. You can disable instant personalization on all pre-approved websites and applications using your Applications and Websites privacy setting. You can also block a particular pre-approved website or application by clicking "No Thanks" in the blue bar when you visit that application or website. In addition, if you log out of Facebook before visiting a pre-approved application or website, it will not be able to access your information.

Exporting Information. You (and those you make your information available to) may use tools like RSS feeds, mobile phone address book applications, or copy and paste functions, to capture, export (and in some cases, import) information from Facebook, including your information and information about you. For example, if you share your phone number with your friends, they may use third party applications to sync that information with the address book on their mobile phone.

Advertisements. Sometimes the advertisers who present ads on Facebook use technological methods to measure the effectiveness of their ads and to personalize advertising content. You may opt-out of the placement of cookies by many of these advertisers here. You may also use your browser cookie settings to limit or prevent the placement of cookies by advertising networks. Facebook does not share personally identifiable information with advertisers unless we get your permission.

Links. When you click on links on Facebook you may leave our site. We are not responsible for the privacy practices of other sites, and we encourage you to read their privacy statements.

5. How We Use Your Information

We use the information we collect to try to provide a safe, efficient, and customized experience. Here are some of the details on how we do that:

To manage the service. We use the information we collect to provide our services and features to you, to measure and improve those services and features, and to provide you with customer support. We use the information to prevent potentially illegal activities, and to enforce our Statement of Rights and Responsibilities. We also use a variety of technological systems to detect and address anomalous activity and screen content to prevent abuse such as spam. These efforts may on occasion result in a temporary or permanent suspension or termination of some functions for some users.

To contact you. We may contact you with service-related announcements from time to time. You may opt out of all communications except essential updates on your account notifications page. We may include content you see on Facebook in the emails we send to you.

To serve personalized advertising to you. We don’t share your information with advertisers without your consent. (An example of consent would be if you asked us to provide your shipping address to an advertiser to receive a free sample.) We allow advertisers to choose the characteristics of users who will see their advertisements and we may use any of the non-personally identifiable attributes we have collected (including information you may have decided not to show to other users, such as your birth year or other sensitive personal information or preferences) to select the appropriate audience for those advertisements. For example, we might use your interest in soccer to show you ads for soccer equipment, but we do not tell the soccer equipment company who you are. You can see the criteria advertisers may select by visiting our advertising page. Even though we do not share your information with advertisers without your consent, when you click on or otherwise interact with an advertisement there is a possibility that the advertiser may place a cookie in your browser and note that it meets the criteria they selected.

To serve social ads. We occasionally pair advertisements we serve with relevant information we have about you and your friends to make advertisements more interesting and more tailored to you and your friends. For example, if you connect with your favorite band's page, we may display your name and profile photo next to an advertisement for that page that is displayed to your friends. We only share the personally identifiable information visible in the social ad with the friend who can see the ad. You can opt out of having your information used in social ads on this help page.

To supplement your profile. We may use information about you that we collect from other Facebook users to supplement your profile (such as when you are tagged in a photo or mentioned in a status update). In such cases we generally give you the ability to remove the content (such as allowing you to remove a photo tag of you) or limit its visibility on your profile.

To make suggestions. We use your information, including the addresses you import through our contact importers, to make suggestions to you and other users on Facebook. For example, if another user imports the same email address as you do, we may suggest that you add each other as friends. Similarly, if one of your friends uploads a picture of you, we may suggest that your friend tag you in the picture. We do this by comparing your friend's pictures to information we've put together from the photos you've been tagged in. We may also suggest that you use certain tools and features based on what your friends have used. You can control whether we suggest that another user add you as a friend through your "search for you on Facebook" privacy setting. You can control whether we suggest that another user tag you in a photo by clicking customize from your privacy settings.

To help your friends find you. We allow other users to use contact information they have about you, such as your email address, to find you, including through contact importers and search. You can prevent other users from using your email address to find you using the search section of your privacy settings.

Downloadable Software. Certain downloadable software applications and applets that we offer, such as our browser toolbars and photo uploaders, transmit data to us. We may not make a formal disclosure if we believe our collection of and use of the information is the obvious purpose of the application, such as the fact that we receive photos when you use our photo uploader. If we believe it is not obvious that we are collecting or using such information, we will make a disclosure to you the first time you provide the information to us so that you can decide whether you want to use that feature.

Memorializing Accounts. If we are notified that a user is deceased, we may memorialize the user's account. In such cases we restrict profile access to confirmed friends, and allow friends and family to write on the user's Wall in remembrance. We may close an account if we receive a formal request from the user's next of kin or other proper legal request to do so.

6. How We Share Information

Facebook is about sharing information with others — friends and people in your communities — while providing you with privacy settings that you can use to restrict other users from accessing some of your information. We share your information with third parties when we believe the sharing is permitted by you, reasonably necessary to offer our services, or when legally required to do so. For example:

When you make a payment. When you enter into transactions with others or make payments on Facebook, we will share transaction information with only those third parties necessary to complete the transaction. We will require those third parties to agree to respect the privacy of your information.

When you invite a friend to join. When you ask us to invite a friend to join Facebook, we will send your friend a message on your behalf using your name. The invitation may also contain information about other users your friend might know. We may also send up to two reminders to them in your name. You can see who has accepted your invitations, send reminders, and delete your friends' email addresses on your invite history page. If your friend does not want us to keep their information, we will also remove it at their request by using this help page.

When you choose to share your information with marketers. You may choose to share information with marketers or electronic commerce providers that are not associated with Facebook through on-site offers. This is entirely at your discretion and we will not provide your information to these marketers without your consent.

To help your friends find you. By default, we make certain information you have posted to your profile available in search results on Facebook to help your friends find you. However, you can control who can see some of this information, as well as who can find you in searches, through your privacy settings. We also partner with email and instant messaging providers to help their users identify which of their contacts are Facebook users, so that we can promote Facebook to those users.

To give search engines access to publicly available information. We generally limit search engines' access to our site. We may allow them to access information set to the "everyone" setting (along with your name and profile picture) and your profile information that is visible to everyone. You can change the visibility of some of your profile information using the customize section of your privacy settings. You can also prevent search engines from indexing your profile using the Applications and Websites privacy setting.

To help improve or promote our service. Sometimes we share aggregated information with third parties to help improve or promote our service. But we only do so in such a way that no individual user can be identified or linked to any specific action or information.

To provide you with services. We may provide information to service providers that help us bring you the services we offer. For example, we may use third parties to help host our website,

send out email updates about Facebook, remove repetitive information from our user lists, process payments, or provide search results or links (including sponsored links). These service providers may have access to your personal information for use for a limited time, but when this occurs we implement reasonable contractual and technical protections to limit their use of that information to helping us provide the service.

To advertise our services. We may ask advertisers outside of Facebook to display ads promoting our services. We may ask them to deliver those ads based on the presence of a cookie, but in doing so will not share any other information with the advertiser.

To offer joint services. We may provide services jointly with other companies, such as the classifieds service in the Facebook Marketplace. If you use these services, we may share your information to facilitate that service. However, we will identify the partner and present the joint service provider's privacy policy to you before you use that service.

To respond to legal requests and prevent harm. We may disclose information pursuant to subpoenas, court orders, or other requests (including criminal and civil matters) if we have a good faith belief that the response is required by law. This may include respecting requests from jurisdictions outside of the United States where we have a good faith belief that the response is required by law under the local laws in that jurisdiction, apply to users from that jurisdiction, and are consistent with generally accepted international standards. We may also share information when we have a good faith belief it is necessary to prevent fraud or other illegal activity, to prevent imminent bodily harm, or to protect ourselves and you from people violating our Statement of Rights and Responsibilities. This may include sharing information with other companies, lawyers, courts or other government entities.

Transfer in the Event of Sale or Change of Control. If the ownership of all or substantially all of our business changes, we may transfer your information to the new owner so that the service can continue to operate. In such a case, your information would remain subject to the promises made in any pre-existing Privacy Policy.

7. How You Can Change or Remove Information

Editing your profile. You may change or remove your profile information at any time by going to your profile page and clicking "Edit My Profile." Information will be updated immediately.

Delete uploaded contacts. If you use our contact importer to upload addresses, you can later delete the list on this help page. You can delete the email addresses of friends you have invited to join Facebook on your invite history page.

Deactivating or deleting your account. If you want to stop using your account you may deactivate it or delete it. When you deactivate an account, no user will be able to see it, but it will not be deleted. We save your profile information (connections, photos, etc.) in case you later

decide to reactivate your account. Many users deactivate their accounts for temporary reasons and in doing so are asking us to maintain their information until they return to Facebook. You will still have the ability to reactivate your account and restore your profile in its entirety. When you delete an account, it is permanently deleted from Facebook. You should only delete your account if you are certain you never want to reactivate it. You may deactivate your account on your account settings page or delete your account on this help page.

Limitations on removal. Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy settings, or it was copied or stored by other users. However, your name will no longer be associated with that information on Facebook. (For example, if you post something to another user's profile and then you delete your account that post may remain, but be attributed to an "Anonymous Facebook User.") Additionally, we may retain certain information to prevent identity theft and other misconduct even if deletion has been requested. If you have given third party applications or websites access to your information, they may retain your information to the extent permitted under their terms of service or privacy policies. But they will no longer be able to access the information through our Platform after you disconnect from them.

Backup copies. Removed and deleted information may persist in backup copies for up to 90 days, but will not be available to others.

Non-user contact information. If a user provides your email address to us, and you are not a Facebook user but you want us to delete your address, you can do so on this help page. However, that request will only apply to addresses we have at the time of the request and not to any addresses that users provide to us later.

8. How We Protect Information

We do our best to keep your information secure, but we need your help. For more detailed information about staying safe on Facebook, visit the Facebook Security Page.

Steps we take to keep your information secure. We keep your account information on a secured server behind a firewall. When you enter sensitive information (such as credit card numbers and passwords), we encrypt that information using secure socket layer technology (SSL). We also use automated and social measures to enhance security, such as analyzing account behavior for fraudulent or otherwise anomalous behavior, may limit use of site features in response to possible signs of abuse, may remove inappropriate content or links to illegal content, and may suspend or disable accounts for violations of our Statement of Rights and Responsibilities.

Risks inherent in sharing information. Although we allow you to set privacy options that limit access to your information, please be aware that no security measures are perfect or impenetrable. We cannot control the actions of other users with whom you share your information. We cannot guarantee that only authorized persons will view your information. We cannot ensure that information you share on Facebook will not become publicly available. We are not responsible for third party circumvention of any privacy settings or security measures on Facebook. You can reduce these risks by using common sense security practices such as choosing a strong password, using different passwords for different services, and using up to date antivirus software.

Report Violations. You should report any security violations to us on this help page.

9. Other Terms

Changes. We may change this Privacy Policy pursuant to the procedures outlined in the Facebook Statement of Rights and Responsibilities. Unless stated otherwise, our current privacy policy applies to all information that we have about you and your account. If we make changes to this Privacy Policy we will notify you by publication here and on the Facebook Site Governance Page. If the changes are material, we will provide you additional, prominent notice as appropriate under the circumstances. You can make sure that you receive notice directly by liking the Facebook Site Governance Page.

Top 10 Privacy Settings

Holy Grail of Facebook Privacy

10 Privacy Settings Every Facebook User Should Know

We've updated this guide with the new privacy settings just launched by Facebook. You can get the new Facebook privacy guide now.

Everyday I receive an email from somebody about how their account was hacked, how a friend tagged them in the photo and they want a way to avoid it, as well as a number of other complications related to their privacy on Facebook. Over the weekend one individual contacted me to let me know that he would be removing me as a friend from Facebook because he was "going to make a shift with my Facebook use – going to just mostly family stuff."

Perhaps he was tired of receiving my status updates or perhaps he didn't want me to view photos from his personal life. Whatever the reason for ending our Facebook friendship, I figured that many people would benefit from a thorough overview on how to protect your privacy on Facebook. Below is a step by step process for protecting your privacy.

1. Use Your Friend Lists

I can't tell you how many people are not aware of their friend lists. For those not aware of what friend lists are, Facebook describes them as a feature which allows "you to create private groupings of friends based on your personal preferences. For example, you can create a Friend List for your friends that meet for weekly book club meetings. You can create Friend Lists for all of your organizational needs, allowing you to quickly view friends by type and send messages to your lists."

There are a few very important things to remember about friend lists:

You can add each friend to more than one friend group

Friend groups should be used like "tags" as used elsewhere around the web

Friend Lists can have specific privacy policies applied to them

I'll touch on each of the things listed above in more detail later. A typical setup for groups would be "Friends", "Family", and "Professional". These three groups can then be used to apply different privacy policies. For example, you may want your friends to see photos from the party you were at last night, but you don't want your family or professional contacts to see those photos.

Using friend lists is also extremely useful for organizing your friends if you have a lot of them. For instance I have about 20 friend lists and I categorize people by city (New York, San Francisco, D.C., Tel Aviv, etc), where I met them (conferences, past co-workers, through this blog), and my relationship with them (professional, family, social, etc).

You can configure your friend lists by visiting the friends area of your Facebook.

2. Remove Yourself From Facebook Search Results

My mom is a teacher and one of the first things she asked me when she joined Facebook is how she could make sure her students couldn't see that she was on the site. Understandably my mom doesn't want her middle school students to know what she's up to in her personal life. There are numerous reasons that individuals don't want their information to show up in search results on Facebook, and it's simple to turn off your public visibility.

How to Remove Yourself From Facebook Search Results

Now that you've decided that you would like to remove yourself from Facebook's search results, here's how to do it:

Visit your search privacy settings page

Under “Search Visibility” select “Only Friends” (Remember, doing so will remove you from Facebook search results, so make sure you want to be removed totally. Otherwise, you can select another group, such as “My Networks and Friends” which I believe is the default.)

Click “Save Changes”

By default, Facebook makes your presence visible to the network you are in. Frequently, people aren’t aware of their visibility, so this is one of the first settings that users wish to modify. By selecting “Customize” from the search visibility drop down you can make your settings even more granular.

3. Remove Yourself From Google

Facebook gets A TON of traffic from displaying user profiles in search engines. Not all of your profile is displayed though. Currently the information displayed in the search profile is limited to: your profile picture, a list of your friends, a link to add you as a friend, a link to send you a message, and a list of up to approximately 20 fan pages that you are a member of.

For some people, being displayed in the search engines is a great way to let people get in contact with you, especially if you don’t have an existing website. Facebook also tends to rank high in the search results, so if you want to be easy to find, making your search profile can be a great idea. Many people don’t want any of their information to be public though.

By visiting the same search privacy settings page listed in the previous step, you can control the visibility of your public search listing which is visible to Google and other search engines. You can turn off your public search listing by simply unchecking the box next to the phrase “Create a public search listing for me and submit it for search engine indexing”.

4. Avoid the Infamous Photo/Video Tag Mistake

This is the classic Facebook problem. You let loose for a few hours one night (or day) and photos (or videos) of the moment are suddenly posted for all to view, not just your close friends who shared the moment with you. The result can be devastating. Some have been fired from work after incriminating photos/videos were posted for the boss to see. For others, randomly tagged photos/videos have ended relationships.

At the least, a tagged photo/video can result in personal embarrassment. So how do you prevent the infamous tagged photo or video from showing up in all of your friends news feeds? It’s

pretty simple. First visit your profile privacy page and modify the setting next to “Photos Tagged of You”. Select the option which says “Customize…”.

Select the option “Only Me” and then “None of My Networks” if you would like to keep all tagged photos private. If you’d like to make tagged photos visible to certain users you can choose to add them in the box under the “Some Friends” option. In the box that displays after you select “Some Friends” you can type either individual friends or friend lists.

5. Protect Your Albums

Just because you’ve uploaded photos doesn’t mean that you’ve accurately tagged every photo correctly. This setting is more of a reminder than anything else. Frequently people will turn of their tagged photo visibility to certain friend lists yet keep their photo albums public to the world. If you are trying to make all your photos invisible you must do so on an album by album basis.

There is a specific Photos Privacy page from which you can manually configure the visibility of each album. This is an extremely useful configuration option and I highly recommend that you take advantage of it. This way you can store your photos indefinitely on Facebook yet ensure that the only people that can view your photos are the ones who you really want to see them.

6. Prevent Stories From Showing Up in Your Friends’ News Feeds

Oh, did you really just break up with your girlfriend? I’m sorry to hear that. I’m sure all of your friends and business contacts are also sorry to hear that. I can’t tell you how many awkward relationship status changes I’ve seen. The most regular one I’ve seen recently is when an attractive female ends their relationship and numerous guys hop on the opportunity to console her.

I’ve also seen the end of marriages, as well as weekly relationship status changes as individuals try to determine where their relationship stands with their significant other. My personal policy is to not display a relationship status, but many like to make a public statement out of their relationship. For those individuals, it can be a smart move to hedge against future disasters.

There are a number of ways to control how your relationship status is displayed. The first thing that most people should do is uncheck the box next to “Remove Relationship Status” in the News Feed and Wall Privacy page. In the rare instance that a relationship does uncomfortably end, you can avoid making things more uncomfortable by avoiding a friend notification about it.

Second, your relationship status falls within your “Basic Information” section of your profile. You can control who can see your basic information next to the “Basic Information” setting on

the Profile Privacy page. Keep in mind that other relevant profile information like your gender, birth date, networks, and other settings are visible within your basic information section.

Making your basic information completely invisible to friends probably isn't a good idea, but removing the news feed stories about relationship changes most likely is.

7. Protect Against Published Application Stories

This one is a little more tricky to manage but I'll explain the issue at hand. Frequently when you add an application, a news feed item is immediately published to your profile. One way to get instantly embarrassed is to visit the "Have Sex!" application (found here). This application has no purpose besides telling your friends that you are interested in having sex with them. Without taking any action, the application will post a news feed story to your profile which says the equivalent of "Nick just published to the world that he is having sex!"

This is surely something that none of your professional contacts if any of your contacts are interested in seeing (honestly I'm a bit confused about that application, but that's a different story). That's why it's important to monitor what takes place after you install an application on Facebook. Once you install an application you should visit your profile to ensure that no embarrassing notification has been posted to your profile.

More often then not, nothing will be posted but there are many applications on the platform unfortunately that publish stories without you knowing it. There are two ways to avoid having this happen: don't visit applications or scan your profile every time that you do. Ultimately you shouldn't be concerned about applications that you've built a trusted relationship with but any new applications could potentially post embarrassing notifications.

8. Make Your Contact Information Private

I personally use Facebook for professional and personal use and it can frequently become overwhelming. That's why I've taken the time to outline these ten privacy protection steps. One of the first things I did when I started approving friend requests from people that I hadn't built a strong relationship with, was make my contact information visible only to close contacts.

The contact information is my personal email and phone number. It's a simple thing to set but many people forget to do it. Frequently people we don't know end up contacting us and we have no idea how they got our contact information. Your contact privacy can be edited right from your profile. If you have chosen to enter this information, you should see a "Contact Information" area under the "Info" tab in your profile.

If it displays, you simply click "Edit" and then a screen like the one pictured below will show up.

For each contact item that you have in your profile you should set custom privacy settings (as pictured below) so that contacts that you aren't close to don't have access to your phone number and/or email. It's a small change but it can save you the hassle of being pestered by people you don't know well. Also, protecting your privacy is generally a good practice to get in the habit of doing.

As a side note, this is a great area to take advantage of friend lists. By getting in the habit of grouping your friends, you can ensure that you are navigating Facebook safely through privacy settings that are attached to your friend lists.

9. Avoid Embarrassing Wall Posts

Just because you use Facebook for business doesn't mean your friends do. That's why once in a while a friend of yours will come post something embarrassing or not necessarily "work friendly" and it can end up having adverse effects. That's why Facebook has provided you with the ability to customize your wall postings visibility. You can also control which friends can post on your wall. There are two places you can control these things.

Adjust Wall Posting Visibility

Within your profile page you can control who can view wall postings made by your friends. To do so, click on the "Settings" icon on the wall in your profile page. Next, find the box pictured in the image above and adjust the setting which says "Who can see posts made by friends?" I'd suggest using a strategy similar to the one outlined in the previous step regarding contact information.

Control Who Can Post to Your Wall

In addition to controlling who can view wall postings published by your friends, you also want to control which friends can post on your wall. Not everybody needs to do this, but occasionally you simply want to prevent some people from posting on your page. If you visit the Profile Privacy settings page, there is a section labeled "Wall Posts".

From this area you can completely disable your friends' ability to post on your wall. You can also select specific friend lists that can post on your wall. Personally, I don't really care who can post on my wall but I can understand the need to control who can see those wall postings. If you want to limit who can post wall posts on your profile, this is where you can do it.

10. Keep Your Friendships Private

While it's fun to show off that you have hundreds or thousands of friends on Facebook, some of your friends don't want to live public lives. That's why it's often a good policy to turn off your friends' visibility to others. I've had a number of individuals visit my profile and then selectively pick off friends that are relevant to them for marketing purposes, or other reasons.

Whatever the reason they are doing it, just know that they are … it's part of what makes Facebook so addictive: the voyeuristic nature. Also, your friends are frequently visible to the public through search engines and exposing this information can ultimately present a security risk. To modify the visibility of your friends, visit the Profile Privacy page.

Navigate down to the setting which says "Friends" and then modify the setting to whatever is right for you.

Conclusion

These are just ten ways that you can protect your privacy on Facebook. While there are a few other small things to keep in mind, these ten settings are most important. Keep in mind that while you may have turned off the visibility of many profile sections, there is no way to prevent all photos or videos from being visible if friends of yours make the images visible.

The best way to prevent embarrassing items from showing up on Facebook in the future is to not make bad judgements in your personal life. We're all human though and being completely paranoid about every choice you make is probably not the best way to live your life. Be aware of what privacy settings are available and be conscious of what your friends may be publishing about you.

While you may not want to configure all of the privacy settings outlined, simply knowing how to do so is a great step in the right direction. By following the 10 settings listed above you are well on your way to an embarrassment free future on Facebook!

Consent Form

DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project:	Mapping individual differences in single document manipulation
Investigators:	Dr. Gordon McCalla, Professor, Department of Computer Science (966-4902), mccalla@cs.usask.ca Terry Peckham, Department of Computer Science (966-2666), tep578@mail.usask.ca

This consent should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information is not included here, please contact one of the study investigators listed above. Please take the time to read this form carefully and to understand any accompanying information. Any involvement in this study is purely voluntary and is to be considered an optional activity.

In this study you will be given one half hour to read and then answer questions based upon some learning material. The purpose of this study is to try to gain an understanding about how individual learners differ in how they use online material to answer questions. In addition, you will be asked to fill out a short questionnaire prior to beginning the study to gather background information about yourself and your prior knowledge, if any, of the subject matter that you will be reading and answering questions on. Furthermore, the questionnaire will inquire about your self-reported level of academic performance to help link previous academic achievement with your current level of performance in this study. At the end of this study you will be given more information about the purpose and goals of the study.

The data collected from this study will be used in articles for publication in journals and conference proceedings. Any information collected in this study will be kept confidential and not shared with anyone outside of the research team. Any results that are published are published as a grouped result and not individual result. In cases where an individual result might be published, further consent to publication will be sought from the individual prior to publication using the email address to provided in the study. As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled. This summary will outline the research and discuss our findings and recommendations.

All of the information we collect from you (data logged by the computer, observations made by the experimenters, and your questionnaire responses) will be stored so that your name, student number, nsid, or email address is not associated with it. Any write-ups of the data will not

include any information that can be linked directly to you. The research materials will be stored with complete security throughout the entire investigation.

None of the information collected will be shared with your instructor. None of the information collected will influence your grade in this or other courses.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact one of the investigators listed above.

If you do not participate in this study, the data automatically collected by the learning environment will be discarded and deleted upon your completion of the module.

By choosing the "I Agree" option below, you indicate that you have understood to your satisfaction the information regarding participation in the research project and agree to participate. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Dr. Gordon McCalla, Professor, Dept. Computer Science (306) 966-4902 mccalla@cs.usask.ca
- Research Ethics Office University of Saskatchewan (306) 966-2084
- Office of Applied Research SIAST 1-866-467-4278

Please note that remuneration for this study is a draw for a $20 Tim Hortons Gift Card and is open only to individuals who participate in this study. You are free to opt out at any time and you will still be given the remuneration as stated.

Following your acceptance of this consent form you will be asked to create a user id and password in order to login to the experiment. The system will email you a link to the study after you have created a user id and password. The passwords will be encrypted using a one way hash algorithm (MD5) so that no one will be able to use or figure out your password. When you have finished entering your user id and password a random computer generated id will be created and associated with all of your data so that your name and other information is not directly associated with your data. If you should choose to withdraw from the study at any time, just enter your username and password into the website again and you can choose the option to withdraw at your convenience. You will also be given the option to place yourself on a list to be considered for any future studies. As with your withdrawal from the studies option, you can also remove or add yourself to the list at any time by logging in and checking the appropriate button.

If you would like a copy of this consent form please print it through your browser, or contact one of the investigators listed above. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan and the Applied Research Office at SIAST. If there are any unforeseen developments that arise that may affect your decision to continue to

participation in this study, you will be contacted and advised of this new development so that you can make an informed decision about your involvement in this study.

Date: June 01,2011

☒ I have read this form and by clicking on the button below agree to participate in this experiment.

☒

Appendix 2 Experiment 2

Questions and Rubric

Describe the process of enumeration. (Bloom: 2 Comprehension) (Marzano: 2 Comprehension – Integrating)

The hacker will try to learn the network's topology, what Operating Systems are being used, defensive methods employed by the network, the names and titles of key corporate personal, and where on the network important data is stored. The hacker may use a combination of methods like port scan, whois, ping sweep, dumpster diving and social engineering to accomplish this.

State the two broad categories of Hacking? (Bloom: 1 Knowledge) (Marzano: 1 Retrieval - Recall)

Hacking can be broken into two broad categories: (1) breaking into a site to steal or modify the user's data (2) making the user's site or network unusable.

Define a Phreaker (Bloom 1: Knowledge) (Marzano: 1 – Recall)

A "Phreaker" is a person who hacks into telecommunications services.

Social engineering is an integral part of hacking. Given the various psychological aspects of social engineering discussed, create a high level, overarching description of these processes and how they fit within the hacking process. (Bloom: 5 Synthesis) (Marzano: 3 Analysis – Generalizing)

Documents for Experiment 2A

Information leakage (1)

Crackers love to go "trashing" to find documents that help them piece together the structure of your company, provide clues about what kinds of computer systems you use, and most important, obtain the names, titles, and telephone numbers of your employees. Think for a moment about the documents your company throws out each day and how an attacker could use them. Do your own dumpster dive and see if you find:

• Company phone books;
• Organizational charts;
• Memos;
• Company policy manuals;
• Calendars of meetings, events, and vacations;
• System manuals;
• Printouts of sensitive data or login names and passwords;

• Printouts of source code;
• Disks and tapes;
• Company letterhead and memo forms;
• Outdated hardware (especially hard drives).

These items provide a wealth of information to crackers. A copy of the company phone book is an extremely valuable tool. Knowing who to call and who to impersonate are the first steps to gaining access to sensitive data.

Having the right names and titles at their fingertips can let smart crackers sound as though they actually work for your company. A cracker interested in finding dial-in access numbers will use the phone book to determine the telephone exchange of your company and use a war dialer to find modem phone numbers.

Information leakage (2)

Another common social engineering trick is "shoulder surfing", someone looking over an employee's shoulder while he or she types in a password. Password guessing is an additional easy social engineering technique. If a person can find out personal things about other people, he can usually use that information to guess a password. For example, the names of children, their birthdays and anniversaries or social security number are all likely candidates for guessing as passwords.

Dumpster diving is a messy, but a very successful technique for acquiring trade secrets and other valuable information. No matter how disgusting dumpster diving sounds, it is legal. Once trash is discarded onto a public street or alley, it is considered fair game. "The courts have held that if it is left to be accessed by commercial carters, then it is no longer private property. It is only private property if there is a 'no trespassing' sign and you had to trespass to get into the dumpster. The LAN Times listed the following items as potential security leaks in corporate trash: company phone books, organizational charts, memos, company policy manuals, calendars of meetings, events and vacations, system manuals, printouts of sensitive data or login names and passwords, printouts of source code, disks and tapes, company letterhead and memo forms, and outdated hardware. Trash can provide a rich source of information for any corporate espionage agent. Phone books can give a hacker names and numbers of people to target and impersonate. Organizational charts contain information about people who are in positions of authority within the organization. Memos provide small amounts of useful information for creating authentic looking fake memos. Policy manuals show hackers how secure and insecure a company really is. Calendars can tell an attacker which employees are out of town at a particular time. System manuals, sensitive data, and other sources of technical information can be found in the trash.

Social engineering: the great con game

•Be professional: You don't want someone to not buy what you're doing. You're trying to create an illusion. You're trying to be believable.

•Be calm: look like you belong there.
•Know your mark: Know your enemy. Know exactly how they will react before they do.
•Do not fool a superior scammer: Trying to out scam an observant or smarter person will end in disaster.
•Plan your escape from your scam: Lets say someone is suspicious. Don't burn your bridges and walk away. Save the source.
•Try to be a woman: It's proven that women are more trusted over the phone. Use that to an advantage. Get a woman's help if needed. It's even better if you're actually a woman (a rarity in our biz).
•Watermarks: Learn to make 'em. They are invaluable for a mail scam.
•Business cards and fake names: Use them for professional things.
•Manipulate the less fortunate and the stupid
•Use a team if you have to: Don't be arrogant and overly proud. If you need help, get it.

Social engineering: is it effective?
Social engineering concentrates on the weakest link of the computer security chain. It is often said that the only secure computer is an unplugged one. The fact that you could persuade someone to plug it in and switch it on means that even powered down computers are vulnerable. Also, the human part of a security set-up is the most essential. There is not a computer system on earth that doesn't rely on humans. This means that this security weakness is universal, independent of platform, software, network or equipment age.
Anyone with access to any part of the system, physically or electronically is a potential security risk. Any information that can be gained may be used for social engineering further information. This means even people not considered as part of a security policy can be used to cause a security breach.
Social engineering, once mastered, can be used to gain access on any system despite the platform or the quality of the hardware and software present. It’s the hardest form of attack to defend against because hardware and software alone won’t stop it.
Social engineering has been around as long as man. It can be defined as an outsider tricking legitimate personnel into aiding illicit acts such as supplying proprietary information or allowing inappropriate access. It preys on the weakest link in a security system – the human being. Social engineers are con artists who exploit human vulnerabilities such as ignorance, naiveté and an individual’s natural desire to be liked and helpful.

Social engineering: example 1
At one of CSI's "Meet the Enemy" session several years ago, an attendee challenged a hacker's boast about social engineering, so the hacker gave a live demonstration.
He dialed up a phone company, got transferred around, and reached the Help Desk. "Who's the supervisor on duty tonight?" "Oh, it's Betty." "Let me talk to Betty.“
"Hey Betty, having a bad day?" "No, why?" "Your systems are down." She said, "My systems aren't down, we're running fine." He said, "All of my monitors here are showing that you're

completely offline. Something is really wrong." She said, "I'm not offline." He said, "You better sign off." She signed off. He said, "Now sign on again."
She signed on again. He said, "We didn't even show a blip, we show no change." He said, "Sign off again." She did.
"Betty, I'm going to have to sign on as you here to figure out what happening with your ID. Let me have your user ID and password." So this senior supervisor at the Help Desk tells him her user ID and password.
He said, "I'm signed on as you now and I can't see the difference. Shoot. I know what it is. Let me sign off. Now sign yourself back on again." She did. He said, "I know what it is. You're on day-old files. You think you're online but you're not. You're on day-old files. Do me a favor, what changes all the time? The PIN code. Pull the PIN code file, just read me off the first ten PIN codes you've got there and I'll compare them." She was reading off the first
PIN code when we heard "click."
He said, "I told you I could."

Psychology of social engineering (1)

Trust Relationships - Often times, the social engineer expends time developing a trust relationship with the intended victim, then exploits that trust. Following a series of small interactions with the target that were positive in nature and problem free, the social engineer moves in for the big strike. Chances are the request will be granted.
Desire to be helpful - Social engineers rely on people's desire to be helpful to others. Exploits include asking someone to hold a door, or with help logging on to an account. Social engineers are also aware that many individuals have poor refusal skills, and rely on a lack of assertiveness to gather information.
Guilt - Most individuals attempt to avoid guilt feelings if possible. Social engineers are often masters of psychodrama, creating situations and scenarios designed to tug at heartstrings, manipulate empathy, and create sympathy. If granting the request will lead to avoidance of guilty feelings, the target is more likely to comply. Believing that not granting the requested information will lead to significant problems for the requestor is often enough to weigh the balance in favor of compliance with the request.
Diffusion of responsibility - If the target can be made to believe that they are not solely responsible for their actions, they are more likely to grant the social engineer's request. The social engineer creates situations with many factors that obfuscate and dilute personal responsibility for decision making. The social engineer may drop names of other employees involved in the decision making process, or claim another employee of higher status has authorized the action.
Additional principles not listed in the above article:
Greed - people are always looking for something for nothing. A contest, a bribe, or a gift can motivate people to reveal information that they would normally not reveal.

Conflict – people will often go to extraordinary lengths to avoid conflict. A hysterical user (especially one who claims to have some organizational authority) who demands to have a user name and password reset – or else – will often find the other person complies with their demand.

Psychology of social engineering (2)

Moral duty - This is where an individual complies because they feel it is their moral duty to. Part of this is guilt. People prefer to avoid guilt feelings and so if there is a chance that they will feel guilty they will if possible avoid this outcome.

Additional principles:

Cognitive dissonance - according to Festinger's theory of cognitive dissonance, a person cannot deeply hold two opposite beliefs simultaneously . The tension to resolve these beliefs becomes as strong a drive as the need for food or water. If a victim believes that the hacker is a good person then it will be difficult, perhaps even impossible, for the victim to believe that the hacker is asking for something bad.

Compliant to authority – Most will comply with the request of anyone who has a uniform, badge, or title. A skillful hacker will claim to be a manager, corporate office, police detective or any other authority figure that the situation demands.

Two types of hacking

• Hack type 1: breaking in

– Most complex type of hacking

– Six steps

• Hack type 2: attacking (just plain breaking)

– Viruses

– Denial of service (DOS)

Two types of hacking

Hacking can be broken into two broad categories: (1) breaking into a site to steal or modify the user's data (2) making the user's site or network unusable.

Breaking into a site is a generally more complex task, and requires more skill, than simply destroying a network.

In the following slides we examine the six steps generally used by hackers to break into, and sometimes take over, a computer network.

On the other hand, breaking (tearing down) a network, is relatively easy. Viruses are easy to write and new viruses are appearing at a phenomenal rate. Additionally, the widespread

adoption of broadband by residential users who often have no firewall protection has made distributed denial of service attacks fairly simple for hackers to deploy.

The six degrees of breaking in (1-3)

1. Case the network: discover network topology, hardware, and software

- Passive techniques

- Active techniques

2. Gain a foothold

- Gain access to any obscure location or process

3. Elevate authority

- Root

- Administrator

The six degrees of breaking in (1-3)

Just like a thief examines a building before breaking into it, a skillful hacker examines the target network before breaking into it. The hacker will try to learn the network's topology, what Operating Systems are being used, defensive methods employed by the network, the names and titles of key corporate personal, and where on the network important data is stored. The hacker may use a combination of methods like port scan, whois, ping sweep, dumpster diving and social engineering to accomplish this. This process is often called enumeration.

The hacker's next task is to gain a foothold in the network. A foothold is an entrance into any part of the network, often with only user-level privileges. This might be accomplished with buffer overflows, brute force password attacks, social engineering, or a combination of these techniques.

Once the network has been breached, the hacker then uses techniques to gain root/administrator authority and real control over the network. Techniques like password cracking and exploiting known OS weaknesses may be used during this process. Root/Administrator authority will give the hacker access to all files, all network passwords, and the operating system itself.

The six degrees of breaking in (4-6)

4. Steal or modify data

- Download sensitive information

- Destroy data and programs

5. Make a back door

- Create rogue user account

- Modify/Add Startup files

6. Cover your tracks

The six degrees of breaking in (4-6)

After attaining root/admin status, the hacker is finally in control and able to do some real damage. Files can be stolen, destroyed, or changed. System registry data may be altered, changing the very operation of the network. Now the hacker begins to make plans to be around for a while. Effective hackers assume that the corporate network administrators will eventually find how they first breached the network and close the hole. The solution for the hacker is to create an innocent looking account that can later be used to get back into the network. Alternately, the hacker may insert a hidden program, which will run at startup, granting the hacker access. Planting remote-control software, or replacing common system applications with Trojans, is another common trick.

Finally, the hacker may try to modify the event logs and other files that the system administrator may use to detect the hacker's presence.

What is Hacking?

"Hackers have infiltrated the power grid." "Hackers stole credit card data." "Hackers created a virus." We've all heard and read these stories on the news. What are hackers though? In this article, I will discuss what hacking is, and what hacking isn't.

What is hacking?

While there are many definitions of hacking, a general definition is to modify something to make it work for you. For computers, hacking includes fixing programs until they work. Also, hacking includes modifying the computer hardware to make it work better or tuned to the person's wishes. The type of hacking that the media discusses includes breaking into secure systems to determine their weaknesses and to explore them. However, the media only points out the malicious uses for breaking into systems.

Black Hats, White Hats, Crackers and Phreakers

What do these terms have in common? They're all terms used by hackers to describe hackers. Just like in typical cowboy fashion, the "white hats" are the good guys and the "black hats" are the bad guys. Although the lines are blurred greatly when it comes to hacking. The "White hats" are security experts who try to find the vulnerabilities in programs and systems, and report them to the manufacturers. They would be considered "ethical hackers" because they either have

authorization to break into the system or program, or they do so with the intent of assisting the manufacturer in securing them. The "Black Hats" are the ones who are trying to find those same vulnerabilities and exploit them. "Cracker" is another term for the Black Hat hackers, usually referring to the creation of software cracks to bypass anti-piracy methods. A "Phreaker" is a person who hacks into telecommunications services.

Final thoughts

While the media has portrayed all hackers as malicious and evil, the reality is that some (if not most) of them are either working as/for security professionals or are only hacking to improve their personal experience. Some of the white hats only hack their own systems in order to tweak them to the fullest extent that they can. Most of the "white hat" hackers are working behind the scenes or in the shadows. The media hardly, if ever, discusses them or their work. Through movies, and sound-bites, the media has jumped on the "hackers are bad" bandwagon-- totally overlooking the people who are trying to make their (and our) lives better through their hacking.

Experiment 2B

Questions and Rubric

Discuss 3 of advantages of Nmap. (Bloom 2: Comprehension) (Marzano: 2 Comprehension – Integrating)

Flexible: Supports dozens of advanced techniques for mapping out networks filled with IP filters, firewalls, routers, and other obstacles. This includes many port scanning mechanisms (both TCP & UDP), OS detection, pings sweeps, and more. See the documentation page.
- **Powerful**: Nmap has been used to scan huge networks of literally hundreds of thousands of machines.
- **Portable**: Most operating systems are supported, including Linux, Open/Free/Net BSD, Solaris, IRIX, Mac OS X, HP-UX, Sun OS, and more.
- **Easy**: While Nmap offers a rich set of advanced features for power users, you can start out as simply as "nmap -O -sS *targethost*". Both traditional command line and graphical (GUI) versions are available to suit your preference. Binaries are available for those who do not wish to compile Nmap from source.
- **Free**: The primary goals of the Nmap Project is to help make the Internet a little more secure and to provide administrators/auditors/hackers with an advanced tool for exploring their networks. Nmap is available for free download, and also comes with full source code that you may modify and redistribute under the terms of the GNU General Public License (GPL).
- **Well Documented**: Significant effort has been put into comprehensive and up-to-date man pages, whitepapers, and tutorials.

Explain the “Quest for Root” and relate how you would go about getting it? (Bloom 3 Application) (Marzano: 4 Knowledge Utilization – problem solving)

Attackers will usually attempt to gain access to a system at any user privilege. Lower privileges are often easier to breach than more powerful accounts. Once access has been gained, the hacker will audit the system to find other active accounts. Accounts have many different levels of privileges on the system. The root account in UNIX and the Administrator account in Windows have the highest privileges. The hacker’s goal is to escalate their privilege until they reach root/administrator!

Devise a plan using the steps, tools and techniques you could use if you were to attempt to gain root over a network. (Bloom 5: Synthesis) (Marzano: 4 Knowledge Utilization – Problem solving)

Devise a plan using the steps, tools and techniques you could use if you were to attempt to gain root over a network. (Bloom 5: Synthesis) (Marzano: 4 Knowledge Utilization – Problem solving)

One of the first things you need to do is enumerate the network (port scan, dns, dig, netscan). Once you have an idea about the network topology social engineering will often provide you with an in through a user level id given the information you learn from your enumeration.

If you cannot achieve access through social engineering then a more direct attack vector must take place where you will try to employ various know security flaws based upon what you have learned from your enumeration process.

Once you have gained access (usually to a low level account) auditing the different accounts visible to find out other active accounts. Once you have the other accounts making use of password programs to guess the passwords of other accounts is often a successful venture. This can be done by using default passwords to installed software all the way to cracking passwords. If you cannot gain passwords you could also install a keystroke logger and hope that someone else logs into the computer.

Outline how the operating system on a targeted computer can affect the type of hacking that is performed. (Bloom 4: Analysis) (Marzano: 3 Analysis – Specifying)

Each operating system has its own set of vulnerabilities. If a hacker does not know the operating system of the targeted computer, they are at a great disadvantage since many of the exploits are OS dependent. Often they will start with banner grabbing hoping that the return provides insight into the OS. From there they will again enumerate the computer and try specific attack vectors that are OS dependent. These types of attacks include buffer overflows, NetBIOS, and null sessions to name a few.

Experiment 2B Documents

Casing a network electronically

It is easy to think of electronic casing as the next logical step to use after using social engineering to learn about a network. However, electronic casing (sometimes called scanning or enumeration) may be done before, after or concurrent with social engineering. What can a hacker learn using electronic casing? Many of the same things that can be learned with social engineering. Information like:
- The name of the company that owns a network
- The name of the person(s) in charge of the network
- The topology of the target network.
- Operating systems employed in the target network.
- The blocks of IP addresses used in the network.
- Identify active IP addresses on a network
- Identify TCP/UDP services running

Sam Spade
Sam spade is an Internet investigative tool that includes the following:
- **Ping** sends a packet of information to a given IP address or Domain name. It verifies that the IP address or domain in the box exists, is operational, and is able to take commands.
- **DNS** - Domain Name Server will return the IP address of a named server or the domain name of an IP address (rDNS)
- **Whois** - gives the InterNIC registration data for a given domain, i.e., who owns that domain.
- **IP Block** - Gives the owner of that block of IP numbers.
- **Dig** - Advanced DNS tool. It retrieves all of the available DNS Resource Records (RR) for a given host or domain.
- **Traceroute** - shows the Internet routing path between your IP address and the IP address in question.

Net Scan Tools
What do NetScanTools do?
• Translate an IP address to a hostname, or vice versa.
• Use NSLOOKUP from a graphical interface. Access any name server for DNS records.
• Find the Authoritative DNS for a domain.
• List all the computers registered in a domain.
• Probe Ports on a target computer(s) for TCP services.
• Sweep an IP address range looking for active computers.
• Diagnose network connectivity problems with Ping, Traceroute, TCP Term, and Echo.
• Plot Ping and Traceroute time response graphs.
• Check to see if a domain name has been used with the Whois utility.

• Find the responsible business or persons for a domain.
• Synchronize your computer clock to accurate network time servers.
• Use TCP Term to test text-based services on both standard and non-standard TCP ports.
• View hidden headers on web pages.
• View NetBIOS shares on your local network.
• Gather IP and MAC addresses of shared computers your local network.

NMAP

Nmap is a utility for port scanning large networks, although it works fine for single hosts. The guiding philosophy for the creation of nmap was TMTOWTDI (There's More Than One Way To Do It). This is the Perl slogan, but it is equally applicable to scanners. Sometimes you need speed, other times you may need stealth. In some cases, bypassing firewalls may be required. Not to mention the fact that you may want to scan different protocols (UDP, TCP, ICMP, etc.). You just can't do all this with one scanning mode. And you don't want to have 10 different scanners around, all with different interfaces and capabilities. Thus virtually every scanning technique is incorporated into nmap.

Nmap ("Network Mapper") is an open source utility for network exploration or security auditing. It was designed to rapidly scan large networks, although it works fine against single hosts. Nmap uses raw IP packets in novel ways to determine what hosts are available on the network, what services (ports) they are offering, what operating system (and OS version) they are running, what type of packet filters/firewalls are in use, and dozens of other characteristics. Nmap runs on most types of computers, and both console and graphical versions are available. Nmap is free software, available with full source code under the terms of the GNU GPL. Nmap is ...

- **Flexible**: Supports dozens of advanced techniques for mapping out networks filled with IP filters, firewalls, routers, and other obstacles. This includes many port scanning mechanisms (both TCP & UDP), OS detection, pings sweeps, and more. See the documentation page.
- **Powerful**: Nmap has been used to scan huge networks of literally hundreds of thousands of machines.
- **Portable**: Most operating systems are supported, including Linux, Open/Free/Net BSD, Solaris, IRIX, Mac OS X, HP-UX, Sun OS, and more. Windows support is in beta and we are not distributing binaries yet. See the portability page.
- **Easy**: While Nmap offers a rich set of advanced features for power users, you can start out as simply as "nmap -O -sS *targethost*". Both traditional command line and graphical (GUI) versions are available to suit your preference. Binaries are available for those who do not wish to compile Nmap from source.
- **Free**: The primary goals of the Nmap Project is to help make the Internet a little more secure and to provide administrators/auditors/hackers with an advanced tool for exploring their networks. Nmap is available for free download, and also comes with full source code that you may modify and redistribute under the terms of the GNU General Public License (GPL).

- **Well Documented**: Significant effort has been put into comprehensive and up-to-date man pages, whitepapers, and tutorials.

Exploiting OS vulnerabilities

Banner Grabbing

When a hacker tries to connect to a remote system, even a rejection can bring welcome information. Most telnet, email, FTP, and Web servers will identify not only themselves, but also the operating system on which they are running. For example, attempting to connect to a telnet server might look like this: boson% telnet 192.9.200.111

Trying 192.9.200.11...

SunOS 4.1.3 login:

Clearly identifying the operating system as Sun version 4.1.3

Buffer Overflows

Some operating systems (or programs that interface with users on the web) have specific buffer overflow vulnerabilities that can be exploited by hackers.

NetBIOS

Windows relies heavily on the NetBIOS naming service. Tools like nbtstat, nbtscan, and net view can be used by a hacker to enumerate all the computers in a Windows domain.

Null sessions

Windows 2000 has a vulnerability that can allow a hacker to log on to port 139 as an anonymous users with a null character for a password. This unauthenticated user can now access network information, user names, group names, and even registry keys – all through TCP port 139. This Achilles heel is called a *null session*.

Top 5 software vulnerabilities

Buffer overflows are covered later.

Homegrown crypto is a very complex subject beyond the expertise of most programmers. Poorly written crypto is easily cracked.

Race conditions - This quote from *Hacking: The Basics* by Zachary Wilson:

"Most systems today are "multitasking/multithreaded". This means that can execute more than one program at a time. There is a danger if two programs need to access the same data at the same time. Imagine two programs, ABC and XYZ, each program attempts to modify the same file. In order to modify a file, each program must first read the file into memory, change the contents in memory, then copy the memory back out into the file. The race condition occurs when program ABC reads the file into memory and then makes the change. However, before ABC gets to write the file, program XYZ steps in and does the full read/modify/write on the file. Now program ABC writes their copies back out to the file. Since program ABC started with a copy before XYZ made its changes, all of XYZ's changes will be lost. Since you need to get the sequence of events in just the right order, race conditions are very rare. Attackers usually attempt such actions thousands of times before they get it right, and gain access to the system."

No trust management. Many complex programs assume that when an input is requested from a non-human source
(i.e. another process in the program or another program al together) that the input can be trusted as non-malicious.
Hackers often exploit this vulnerability.
Confusing pseudorandom with random. When a program calls for a random number (i.e. to generate the initial
SEQ number in a TCP connection) the computer generates a pseudorandom random. Because this is not a genuinely
random number, hacker may be able to guess this number. This principle is at the heart of some Man-In-The-
Middle attacks.

Buffer overflows

What is a buffer overflow?

A buffer is a place in the computer's memory where data is temporarily stored until the computer is ready to process the data. The program in the computer's memory allocates the buffer size (in other words, the size is primarily dependent on software – not hardware). Most programmers do not expect users to input more data than required. For example, when a programmer prompts the user with: **Please enter your birth date (MM/DD/YY)** the programmer expects an input like: **4/12/62**.

Most programmers do not expect: **232ljl54562323mkjfk…ds9ud0s9fjlkio. 9r98r92. nkfkl. djkfjd999999kmmwq43h–0w wposre-=0o4,<<<yc.-4%*arxxsp+++12,.$)@.** What does the computer do with this extra data? In a well-written program, the computer will not accept more data than has been allocated for the buffer. In a poorly designed program, this data is accepted and pushes past the allocated buffer space into an adjacent part of the memory. This "extra data" could crash the program. In the hands of a skilled hacker, this "extra data" will contain new instructions for the computer allowing the hacker to take over the system. Buffer overflows remain a persistent and pernicious problem in the computers and networking equipment.

The Quest for Root

Attackers will usually attempt to gain access to a system at any user privilege. Lower privileges are often easier to breach than more powerful accounts. Once access has been gained, the hacker will audit the system to find other active accounts. Accounts have many different levels of privileges on the system. The root account in UNIX and the Administrator account in Windows have the highest privileges. The hacker's goal is to escalate their privilege until they reach root/administrator!

Default Passwords

Many vendors have a dirty little secret: they ship software and hardware with default usernames and passwords, some of which they do not tell customers about. Once an attacker knows these default settings they can typically access the software remotely and gain administrative control. This can be extremely dangerous. Consider an attacker gaining access over your switch and

routing infrastructure and forwarding traffic from the R&D department to another server. Alternatively, imagine the attacker taking over your remote access devices, such as ISDN routers, and then sniffing passwords as users access the corporate LAN. This is a huge problem because companies buy lots and lots of hardware and software that they need to deploy quickly. This often results in minimal configuration effort being made, and the default passwords are usually left in, due to carelessness, or for the simple fact that the people installing it don't know (hardware vendors like 3Com have placed backdoors in hardware so that they can help the customer recover)

Password Cracking

This site concerns the practical demonstration of cryptography weakness. (If you don't like the term "password cracker", use "password recovery" instead). All software presented here illustrates four main reasons of cryptosystems untrustworthiness: application of weak algorithms, wrong implementation or application of crypt algorithms and human factor. The main goal of this site is to explain these reasons and to convince people of the use of strong cryptography. Therefore you can't find here password "crackers" that do not use any weakness, such as http, POP3 or SMTP password crackers, Hotmail password cracker, etc. Only free password crackers or best of commercial ones are included.

Key stroke logger (Trojan)

There are several programs going around that make any virus you have seen to date seem like harmless child's play.

These programs will allow anyone on the Internet to remotely control your computer. They can collect all your passwords, access all your accounts including Email and PeopleSoft, read and modify all your documents, publish your hard drive so its shared across the Internet, record your keystrokes, look at your screen, and listen to your conversations on your computer's microphone. You'll never know its happening.

Consider for a moment the implications of someone controlling your computer. They would have access to any resources you access from your computer. If you access your employer's systems, they could use those accounts to perform fraudulent transactions. They could perform online stock or banking transactions with your personal accounts. They could read your email and send email in your name. They could use your computer as a stepping stone to another computer in which case you may get blamed. The victims of any abuse performed by the controller of your computer would only see your computer's network address. You may even be sitting in front of the keyboard when the computer is used in some crime. This would make it very difficult for you to prove your innocence, particularly if the actual perpetrator erased the evidence of their presence after performing the crime.

The programs have been disguised as games, pictures, screen savers, holiday greetings, and other files. The three most popular are probably Netbus, Back Orifice, and SubSeven. However, there are ***hundreds*** more. We'll refer to all of them here as Remote Control Trojan Horse (RCTH)

Programs. They can be used by anyone more sophisticated than a precocious ten year old to compromise your computer.

Network Attacks

Packet sniffiers

Packet sniffiers use the Ethernet card in promiscuous mode to monitor all network traffic on a segment. Think of this like a wiretap on a telephone, except that dozens of conversations may be monitored simultaneously. Some sniffer programs, such as tcpdump, capture only raw data, while others, such as sniffit, have built-in packet reconstruction algorithms and address filters. Some may have an explicit password filter, producing a compact log of username/password/host data. Some even send detected passwords to an IRC channel in realtime.

SYN DoS & Ping of death

The TCP SYN Attack

When a normal TCP connection starts, a destination host receives a SYN (synchronize/start) packet from a source host and sends back a SYN ACK (synchronize acknowledge). The destination host must then hear an ACK (acknowledge) of the SYN ACK before the connection is established. This is referred to as the "TCP three-way handshake." While waiting for the ACK to the SYN ACK, a connection queue of finite size on the destination host keeps track of connections waiting to be completed. This queue typically empties quickly since the ACK is expected to arrive a few milliseconds after the SYN ACK. The TCP SYN attack exploits this design by having an attacking source host generate TCP SYN packets with random source addresses toward a victim host. The victim destination host sends a SYN ACK back to the random source address and adds an entry to the connection queue. Since the SYN ACK is destined for an incorrect or nonexistent host, the last part of the "three-way handshake" is never completed and the entry remains in the connection queue until a timer expires, typically for about one minute. By generating phony TCP SYN packets from random IP addresses at a rapid rate, it is possible to fill up the connection queue and deny TCP services (such as e-mail, file transfer, or WWW) to legitimate users. There is no easy way to trace the originator of the attack because the IP address of the source is forged. The external manifestations of the problem include inability to get e-mail, inability to accept connections to WWW or FTP services, or a large number of TCP connections on your host in the state SYN_RCVD.

Distributed Denial of Service

In a DDoS attack, the attacking packets come from tens or hundreds of addresses rather than just one, as in a "standard" DoS attack. Any DoS defense that is based upon monitoring the volume of packets coming from a single address or single network will then fail since the attacks come from all over. Rather than receiving, for example, a thousand gigantic Pings per second from an attacking site, the victim might receive one Ping per second from 1000 attacking sites. One of the other disconcerting things about DDoS attacks is that the handler can choose the location of the agents. So, for example, a handler could target several NATO sites as victims and employ agents

that are all in countries know to be hostile in NATO. The human attacker, of course, might be sitting in Canada. Like DoS attacks, all of the DDoS attacks employ standard TCP/IP messages -- but employ them is some non-standard ways. Common DDoS attacks have such names as Tribe Flood Network (TFN), Trin00, Stacheldraht, and Trinity. Some details about these will be presented in the following sections.

Distributed Denial of Service

How does the hacker install the DDoS software on the zombie systems?

A typical installation might go something like this:

1). A stolen account is set up as a repository for pre-compiled versions of scanning tools, attack (i.e. buffer overrun exploit) tools, root kits and sniffers, trinoo daemon and master programs, lists of vulnerable hosts and previously compromised hosts, etc. This would normally be a large system with many users, one with little administrative oversight, and on a high-bandwidth connection for rapid file transfer.

2). A scan is performed on large ranges of network blocks to identify potential targets. Targets would include systems running various services known to have remotely exploitable buffer overflow security bugs … stolen accounts on any architecture can be used for caching tools and log files.

3). A list of vulnerable systems is then used to create a script that performs the exploit, sets up a command shell running under the root account that listens on a TCP port (commonly 1524/tcp, the "ingreslock" service port), and connects to this port to confirm the success of the exploit. In some cases, an electronic mail message is sent to an account at a free web based email service to confirm which systems have been compromised. The result is a list of "owned" systems ready for setting up back doors, sniffers, or the trinoo daemons or masters.

Smurf attack

In the "smurf" attack, attackers are using ICMP echo request packets directed to IP broadcast addresses from remote locations to generate denial-of-service attacks. There are three parties in these attacks: the attacker, the intermediary
(aka zombies), and the victim (note that the intermediary can also be a victim).

The intermediary receives an ICMP echo request packet directed to the IP broadcast address of their network. If the intermediary does not filter ICMP traffic directed to IP broadcast addresses, many of the machines on the network will receive this ICMP echo request packet and send an ICMP echo reply packet back. When (potentially) all the machines on a network respond to this ICMP echo request, the result can be severe network congestion or outages.

When the attackers create these packets, they do not use the IP address of their own machine as the source address.

Instead, they create forged packets that contain the spoofed source address of the attacker's intended victim. The result is that when all the machines at the intermediary's site respond to the ICMP echo requests, they send replies to the victim's machine. The victim is subjected to network congestion that could potentially make the network unusable. Even though we have not

labeled the intermediary as a "victim," the intermediary can be victimized by suffering the same types of problem that the "victim" does in these attacks.

Attackers have developed automated tools that enable them to send these attacks to multiple intermediaries at the same time, causing all of the intermediaries to direct their responses to the same victim. Attackers have also developed tools to look for network routers that do not filter broadcast traffic and networks where multiple hosts respond. These networks can the subsequently be used as intermediaries in attacks.

Consent Form

DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project:	Mapping individual differences in single document manipulation
Investigators:	Dr. Gordon McCalla, Professor, Department of Computer Science (966-4902), mccalla@cs.usask.ca Terry Peckham, Department of Computer Science (966-2666), tep578@mail.usask.ca

This consent should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information is not included here, please contact one of the study investigators listed above. Please take the time to read this form carefully and to understand any accompanying information. Any involvement in this study is purely voluntary and is to be considered an optional activity.

In this study you will be given one half hour to read and then answer questions based upon some learning material. The purpose of this study is to try to gain an understanding about how individual learners differ in how they use online material to answer questions. In addition, you will be asked to fill out a short questionnaire prior to beginning the study to gather background information about yourself and your prior knowledge, if any, of the subject matter that you will be reading and answering questions on. Furthermore, the questionnaire will inquire about your self-reported level of academic performance to help link previous academic achievement with your current level of performance in this study. At the end of this study you will be given more information about the purpose and goals of the study.

The data collected from this study will be used in articles for publication in journals and conference proceedings. Any information collected in this study will be kept confidential and not shared with anyone outside of the research team. Any results that are published are published as a grouped result and not individual result. In cases where an individual result might be published, further consent to publication will be sought from the individual prior to publication using the email address to provided in the study. As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled. This summary will outline the research and discuss our findings and recommendations.

All of the information we collect from you (data logged by the computer, observations made by the experimenters, and your questionnaire responses) will be stored so that your name, student number, nsid, or email address is not associated with it. Any write-ups of the data will not

include any information that can be linked directly to you. The research materials will be stored with complete security throughout the entire investigation.

None of the information collected will be shared with your instructor. None of the information collected will influence your grade in this or other courses.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact one of the investigators listed above.

If you do not participate in this study, the data automatically collected by the learning environment will be discarded and deleted upon your completion of the module.

By choosing the "I Agree" option below, you indicate that you have understood to your satisfaction the information regarding participation in the research project and agree to participate. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Dr. Gordon McCalla, Professor, Dept. Computer Science (306) 966-4902 mccalla@cs.usask.ca
- Research Ethics Office University of Saskatchewan (306) 966-2084
- Office of Applied Research SIAST 1-866-467-4278

Please note that remuneration for this study is a draw for a $20 Tim Hortons Gift Card and is open only to individuals who participate in this study. You are free to opt out at any time and you will still be given the remuneration as stated.

Following your acceptance of this consent form you will be asked to create a user id and password in order to login to the experiment. The system will email you a link to the study after you have created a user id and password. The passwords will be encrypted using a one way hash algorithm (MD5) so that no one will be able to use or figure out your password. When you have finished entering your user id and password a random computer generated id will be created and associated with all of your data so that your name and other information is not directly associated with your data. If you should choose to withdraw from the study at any time, just enter your username and password into the website again and you can choose the option to withdraw at your convenience. You will also be given the option to place yourself on a list to be considered for any future studies. As with your withdrawal from the studies option, you can also remove or add yourself to the list at any time by logging in and checking the appropriate button.

If you would like a copy of this consent form please print it through your browser, or contact one of the investigators listed above. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan and the Applied Research Office at SIAST. If there are any unforeseen developments that arise that may affect your decision to continue to

participation in this study, you will be contacted and advised of this new development so that you can make an informed decision about your involvement in this study.

Date: June 01,2011

☒ I have read this form and by clicking on the button below agree to participate in this experiment.

☒Bottom of Form

Experiment 2 Statistics

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	1.807	3	0.602	42.962	3.86E-9	88.67
within groups	0.294	21	0.014			11.33
total	2.101	24				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.1767	0.5182	0.5518	0.866154	-	-
Gabriel comparison interval	0.108	0.171	0.108	0.067	-	-
n	5	2	5	13	-	-

Statistics for Bloom Level 1

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.27608	0.2087	0.17365
60,30,10	0.3415*	-	0.27608	0.25064
70,20,10	0.3751*	0.0336	-	0.17365
80,10,10	0.6895*	0.348*	0.31435*	-

Tukey-Kramer for Bloom Level 1

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	0.628	2	0.314	51.904	2.00E-8	91.43
within groups	0.115	19	0.00605			8.57
total	0.743	21				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.335967	-	0.633767	0.821688	-	-
Gabriel comparison interval	0.083	-	0.083	0.036	-	-
n	3	-	3	16	-	-

Statistics for Bloom Level 2

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0.16135	0.12433
60,30,10				
70,20,10	0.2978*		-	0.12433
80,10,10	0.4857*		0.18792*	-

Tukey-Kramer for Bloom Level 2

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	0.598	2	0.299	11.117	0.023	82.52
within groups	0.108	4	0.027			17.48
total	0.706	6				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.239033	-	0.4284	0.863733	-	-
Gabriel comparison interval	0.00E0	-	0.00E0	0.00E0	-	-
n	3	-	1	3	-	-

Statistics for Bloom Level 3

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0	0
60,30,10				
70,20,10	0.18937*		-	0
80,10,10	0.6247*		0.4353*	-

Tukey-Kramer for Bloom Level 3

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	0.371	1	0.371	15.434	0.017	82.79
within groups	0.096	4	0.024			17.21
total	0.467	5				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.3187	-	-	0.815867	-	-
Gabriel comparison interval	0.00E0	-	-	0.00E0	-	-
n	3	-	-	3	-	-

Statistics for Bloom Level 4

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-			0
60,30,10				
70,20,10				
80,10,10	0.4972*			-

Tukey-Kramer for Bloom Level 4

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	6.356	3	2.119	169.753	2.41E-34	90.99
within groups	0.986	79	0.012			9.01
total	7.342	82				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.222957	0.5182	0.566277	0.864164	-	-
Gabriel comparison interval	0.044	0.151	0.059	0.032	-	-
n	23	2	13	45	-	-

Statistics for Bloom Level 5

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.21617	0.10175	0.07516
60,30,10	0.29524*	-	0.22272	0.2119
70,20,10	0.3433*	0.04808	-	0.09233
80,10,10	0.6412*	0.346*	0.29789*	-

Tukey-Kramer for Bloom Level 5

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	3.505	3	1.168	120.982	3.32E-27	87.32
within groups	0.657	68	0.009657			12.68
total	4.162	71				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.386605	0.64298	0.749081	0.934512	-	-
Gabriel comparison interval	0.041	0.059	0.047	0.038	-	-
n	21	10	16	25	-	-

Statistics for Main Marzano Level 1

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.09944	0.08589	0.07661
60,30,10	0.25638*	-	0.10434	0.09684
70,20,10	0.3625*	0.1061*	-	0.08286
80,10,10	0.5479*	0.29153*	0.18543*	-

Tukey-Kramer for Marzano Level 1

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	0.978	3	0.326	62.312	1.27E-10	91.63
within groups	0.11	21	0.005232			8.37
total	1.088	24				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.469255	0.6686	0.75295	0.921975	-	-
Gabriel comparison interval	0.044	0.104	0.074	0.052	-	-
n	11	2	4	8	-	-

Statistics for Main Marzano Level 2

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.15497	0.11771	0.09368
60,30,10	0.19935*	-	0.17459	0.15938
70,20,10	0.2837*	0.08435	-	0.12345
80,10,10	0.4527*	0.25338*	0.16903*	-

Tukey-Kramer for Marzano Level 2

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	0.821	3	0.274	52.718	8.05E-9	92.50
within groups	0.088	17	0.00519			7.50
total	0.909	20				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.3517	0.6083	0.72856	0.964642	-	-
Gabriel comparison interval	0.106	0.106	0.067	0.043	-	-
n	2	2	5	12	-	-

Statistics for Main Marzano Level 3

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.20479	0.17134	0.15641
60,30,10	0.2566*	-	0.17134	0.15641
70,20,10	0.3769*	0.12026	-	0.10901
80,10,10	0.6129*	0.3563*	0.23608*	-

Tukey-Kramer for Marzano Level 3

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	470.581	3	156.86	0.604	0.626	-
within groups	2858.367	11	259.852			-
total	3328.948	14				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.00E00	12.65192	0.46995	0.985871	-	-
Gabriel comparison interval	35.817	16.018	25.327	13.538	-	-
n	1	5	2	7	-	-

Statistics for Main Marzano Level 4

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	53.157	59.4313	51.8759
60,30,10	12.652	-	40.5993	28.4136
70,20,10	0.47	12.182	-	38.9069
80,10,10	0.9859	11.666	0.5159	-

Tukey-Kramer for Marzano Level 4

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	2.351	3	0.784	43.863	6.77E-10	86.67
within groups	0.429	24	0.018			13.33
total	2.779	27				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.19575	0.67886	0.50394	0.93905	-	-
Gabriel comparison interval	0.11	0.12	0.12	0.078	-	-
n	6	5	5	12	-	-

Statistics for Main Marzano Combined with Previous Data Level 4

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.22324	0.22324	0.18434
60,30,10	0.4831*	-	0.23317	0.19624
70,20,10	0.30819*	0.17492	-	0.19624
80,10,10	0.7433*	0.26019*	0.4351*	-

Tukey-Kramer for Main Marzano Combined with Previous Data Level 4

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	4.919	10	0.492	31.471	2.54E-24	79.16
within groups	1.313	84	0.016			20.84
total	6.232	94				

	A50,30,20	A60,30,10	A70,20,10	A80,10,10	B50,30,20	B60,30,10	B70,20,10	B80,10,10	C50,30,20	C60,30,10	C70,20,10	C80,10,10
mean	0.33593	0.646475	0.7626	0.890767	0.351977	0.614429	0.688081	0.893219	0.4272	-	0.7465	0.995583
Gabriel comparison interval	0.096	0.151	0.214	0.124	0.084	0.114	0.076	0.058	0.175	-	0.303	0.124
n	10	4	2	6	13	7	16	27	3	-	1	6

Statistics Marzano 1 Sublevels

	A50,30,20	A60,30,10	A70,20,10	A80,10,10	B50,30,20	B60,30,10	B70,20,10	B80,10,10	C50,30,2	C60,30,10	C70,20,10	C80,10,10
A50,30,20	-	0.24477	0.32048	0.213651511	0.174025882	0.203890525	0.16678	0.15316	0.27235		0.433928	0.21365
A60,30,10	0.31055*	-	0.3583	0.267064389	0.236561678	0.259321742	0.23128	0.22166	0.31599		0.462569	0.26706
A70,20,10	0.4267*	0.11613	-	0.3378127	0.314253815	0.331725486	0.3103	0.3032	0.37769		0.506719	0.33781
A80,10,10	0.5548*	0.24429	0.12817	-	0.20419782	0.230180555	0.19806	0.18673	0.29255		0.446884	0.23887
B50,30,20	0.016047	0.2945*	0.4106*	0.5388*	-	0.193961665	0.15449	0.13967	0.265		0.429352	0.2042
B60,30,10	0.2785*	0.03205	0.14817	0.27634*	0.26245*	-	0.18749	0.17548	0.2855		0.442301	0.23018
B70,20,10	0.3522*	0.04161	0.07452	0.20269*	0.3361*	0.07365	-	0.13053	0.2603		0.426468	0.19806
B80,10,10	0.5573*	0.24674*	0.13062	0.0024519	0.5412*	0.27879*	0.20514*	-	0.25179		0.421326	0.18673
C50,30,20	0.09127	0.21928	0.3354	0.4636*	0.07522	0.18723	0.26088*	0.466*	-		0.477739	0.29255
C60,30,10												
C70,20,10	0.4106	0.10003	0.0161	0.14427	0.3945	0.13207	0.05842	0.14672	0.3193		-	0.44688
C80,10,10	0.6597*	0.3491*	0.23298	0.10482	0.6436*	0.3812*	0.3075*	0.10236	0.5684*		0.24908	-

Tukey-Kramer Marzano 1 Sublevels

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	0.753	2	0.376	65.063	1.76E-9	92.24
within groups	0.116	20	0.005785			7.76
total	0.869	22				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.343875	-	0.633767	0.821688	-	-
Gabriel comparison interval	0.07	-	0.081	0.035	-	-
n	4	-	3	16	-	-

Statistics Marzano Sublevel 2.1

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0.14698	0.10758
60,30,10				
70,20,10	0.28989*		-	0.12107
80,10,10	0.4778*		0.18792*	-

Tukey-Kramer Marzano Sublevel 2.1

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)		
among groups	**3.493**	**6**	**0.582**	**57.633**	**4.43E-16**	**91.34**		
within groups	**0.333**	**33**	**0.01**			**8.66**		
total	**3.827**	**39**						
	A50,30,20	A60,30,10	A70,20,10	A80,10,10	B50,30,20	B60,30,10	B70,20,10	0
mean	**0.229186**	**-**	**0.617267**	**0.919992**	**0.2903**	**0.6083**	**0.7259**	**0.969056**
Gabriel comparison interval	**0.088**	**-**	**0.134**	**0.067**	**0.116**	**0.164**	**0.134**	**0.077**
n	7	-	3	12	4	2	3	9

Statistics Marzano Sublevel 3

	A50,30,20	A60,30,10	A70,20,10	A80,10,10	B50,30,20	B60,30,10	B70,20,10	B80,10,10
A50,30,20	-		0.2176	0.14997	0.19764	0.25283	0.2176	0.15891
A60,30,10								
A70,20,10	0.3881*		-	0.20354	0.24084	0.28785	0.25746	0.21022
A80,10,10	0.6908*		0.30273*	-	0.18205	0.24084	0.20354	0.13905
B50,30,20	0.06111		0.327*	0.6297*	-	0.27308	0.24084	0.18949
B60,30,10	0.3791*		0.008967	0.31169*	0.318*	-	0.28785	0.2465
B70,20,10	0.4967*		0.10863	0.19409	0.4356*	0.1176	-	0.21022
B80,10,10	0.7399*		0.3518*	0.04906	0.6788*	0.3608*	0.24316*	-

Tukey-Kramer Marzano Sublevel 3

Appendix 3 Experiment 3

The Journal Question

Read the following question carefully. Write a journal response following the advice we discussed in class. Read the rubric so that you know what is expected. Be sure to write enough, make sure you include ideas from the readings you are given, and include higher levels of thinking in your response. Do NOT summarize the texts although you may include small bits of information that help explain the points you are making in the journal response. Here is the question. It is long. Read it all!!

1. **Comment on the following grand narrative that many people worldwide believe about the Canadian identity:**

   "Canada is a democratic, multicultural country free from racism and violence. Canadian citizens are caring and tolerant people that have a global reputation for peacekeeping."

   This is one of the quotes that Sheelah Mclean likes to discuss with audiences around the world. For those of you who don't know Sheelah Mclean, she is one of the founders of Idle No More, a group that stands up for Aboriginal rights in Canada. **Do you agree or disagree with this quote (grand narrative)? Why or why not?** Be sure to include evidence from the readings that have been given to you.

Q2) When were the hyphenated Canadians beginning to flourish?

Finding a Nationality that Fits - Around 1975 (fifth and sixth paragraphs)

Q3) In John Wayne rides again, who owns the land?

Answer: God own the land (line 7)

Q4) Discuss the thought that money grows on trees in Canada.

Student Sample (First Paragraph) The perception of an outsider is that money grow on trees and machines do all the work for you and everything is easy and all your dreams come true. Contrary to this belief is that you actually have to work hard for your dreams to come true.

Documents

Conflict and Multiculturalism

Many people come to Canada because of conflict in their countries. They may have been at war with another country or with a particular group of people within their own country. Sometimes newcomers have trouble forgetting the hostility they feel towards the other group and may continue to have conflicts with members of that group here in Canada. In a survey conducted in Edmonton, ESL students were asked the following question: "What do you think about people bringing their conflicts to Canada?" Here are some of their comments:

"We live in Canada now. People who were our enemies are not our enemies here."
~ Woman from Iraq

"People come here to get away from war, not to make war."
~ Man from El Salvador

"We should find a middle way; there should be general rules for everyone, and we must respect Canadian law."
~ Man from Romania

"It is bad to bring conflicts from your home country. People should get along even with their enemies here."
~Woman from Bosnia

"Canada is a peaceful country. Immigrants should respect that."
~Woman from Korea

Taken from:
Cameron, J., Derwing, T. (2010). *Being Canadian*, 3rd Ed.. Saint-Laurent, Quebec: Pearson.

Essay: "Finding a Nationality that Fits" by Isabel Vincent

From *Pens of Many Colors*

We started to become Canadian the day my mother got her first pair of pants.

They were gray-green gabardine with a high waist, and came wrapped in tissue paper in an Eaton's box. My mother reluctantly modeled them for my brother and me, all the while declaring that she couldn't imagine ever feeling comfortable with the stretchy cloth hugging her hips. Portuguese women didn't wear pants, only the *canadians* dared wear anything so revealing. But in the same breath she'd rationalize that she spent too much money not to wear them, and besides they'd probably be warm in winter.

That was in 1975, a few years after my family had made the big break and moved from the poor immigrant enclave of Kensington Market to the more upscale neighborhood of North York, where pockets of European immigration were just beginning to emerge. We were pioneers in a way. My father had been among the first wave of Portuguese immigrants to Canada in the early fifties, working a bleak stretch of railroad near Port Arthur – now Thunder Baby, Ont. – to earn enough money for my mother's passage across the Atlantic. My mother arrived sea-sick in Halifax in 1955, and took a slow train to Toronto, where she joined my father in a roach-infested flat on Nassau Avenue in the Market.

My mother still speaks of those early *sacrificios*; living in a cold climate with cockroaches and mutely shopping for groceries, pointing out items to a local shopkeeper because she couldn't speak English. Her language skills were so tenuous that she once interpreted a greeting from an Orthodox Jew who lived in the neighborhood as an offer to buy my brother.

In those days, Toronto police used to disperse small crowds of Portuguese men who lingered too long outside cafes. Despite a burgeoning group of immigrants, there were few Portuguese speakers, even in the market.

But by 1975, the market became a Saturday-morning diversion for us, a place to shop for salted cod and fresh vegetables. To the hearty Portuguese immigrants who still worked in the factories and construction yards, and rented windowless basements in the market, we were on our way up. After all, there were very few Portuguese families north of Eglinton Avenue. Although we lived in a mostly Jewish and Italian neighborhood, we were finally becoming Canadian. Or so I thought.

I learned English in my first year of school. Multiculturalism was just beginning and hyphenated Canadians were beginning to flourish. I played with Italian-Canadians, Lithuanian-Canadians and Chinese-Canadians, but at that time nobody – especially suburban 7-year olds – seemed able to pronounce "Portuguese-Canadian," so I told people I was Greek; it was easier to say. My brother went even further, changing his name to something faintly Anglo-Saxon, so his teachers and classmates wouldn't get tongue-tied around those sloshy Portuguese vowels and embarrass him. It seemed a very practical idea at the time, and I reluctantly followed suit.

But we still had problems, and didn't seem to belong. We never quite fit into the emerging Portuguese community, growing up around the parish of St. Mary's Church and the Toronto branch of the popular Benfica soccer club on Queen Street West. We were strangely

aloof with our compatriots, most of whom had emigrated from the Azores, and whose guttural form of Portuguese we had difficulty understanding. My brother and I balked at heritage-language classes and remained passive spectators at the annual religious processions.

But if we had trouble dealing with our peers in downtown Toronto, in North York we were not much better off. My mother and aunts spoke disparagingly of the *canadianas*, Canadian women who (they were sure) knew nothing about how to keep a clean house or cook a decent meal. My mother taught me to cook and sew, and she and my aunts teased my brother, saying someday he'd marry a *canadiana* and would end up doing all the housework.

For all her predictions, my mother was delighted to find out that she had been wrong. My borther, a physician, did marry a Canadian, but he doesn't do much of the housework. These days, my mother's biggest problem is pronouncing the name of her new grandson, Matthew Loughlin MacLean Vincent.

As I grew older I developed a nostalgia for my Lusitanian past, and tried desperately to reintegrate into the community. But I soon grew to hate the hypocrisy of some of my compatriots, most of whom were immigrants who chose to spend several years working in Canada, only to retire to the Portuguese country-side and build their palatial retreats with the fat pensions they collected from the Canadian government. Like my father, who learned English quickly and severed ties with his homeland, I became a staunch Canadian. I could sing *The Maple Leaf Forever* before I was 10, and spent my childhood years in French immersion. I became so good at masking my heritage that a few years ago when I applied for a job at a Toronto newspaper I was turned down because I was perceived as being too Anglo-Saxon.

"If you were ethnic, I'm sure they would have hired you on the spot," the wife of the paper's managing editor told me a year later.

But for most of my life being Portuguese seemed to me a liability. And then my mother bought that important first pair of pants. For a while it seemed that my life had changed. I was proud of my mother: she was becoming like all of the other mothers in the neighborhood. But my excitement was short-lived. A few days later, she decided they just wouldn't do. She carefully wrapped them back up in the tissue paper, placed them in the cardboard Eaton's box, and returned them to the store.

John Wayne rides again by Richard Wagamese

There's a new anthem making the rounds in Indian country.

It goes:

O Canada, your home's on Native land,

With new patriot love we balk at your command.
With knowing hearts, we seethe and rise
The Mohawk, Blood and Cree.
And guard our stand, O Canada.
We'll guard our stand on thee.
God owns the land,
Not you or me.
O Canada we take a stand on thee.
O Canada we take a stand on thee.

While mainstream acceptance of the revised version may be a trifle slow in coming, the lyrics indicate the growing dissatisfaction on the part of aboriginal groups across the country and the accompanying unity. The times - as the old song goes - they are a-changing.

http://www.ammsa.com/node/17522 (online version)

Reference:

Book: Windspeaker
Author: Richard Wagamese
Volume: 8
Issue: 11
Year: 1990
Page 4

Student Sample

This is a response journal that a student wrote. The story is very effective. She talks about life as a new Canadian. However, she did not have the same question as you have NOR does she include any examples from the text. However, this is helpful in that it gives you an idea of something you might say from your own life.

What does it mean to be a Canadian?

Being a newcomer to Canada, I have a unique perspective on being Canadian. I grew up in Pakistan where Canada was talked about as a wonderful land where dreams come true. Whenever we talk about Canada, we considered Canada as a land of comfort, where money grows on trees, and machines do all the work for you! You don't have to do any work, you just sit and rest and let the machines do everything. In this wonderful land, where machines cook for

you, and machines wash your dishes, I could see my dreams coming true. I would imagine myself arriving in Canada, going to university and fulfilling my dream of becoming a doctor, while machines did work for me, and money just comes to me without working hard.

As soon as I arrived in Canada, I saw a very different country. Yes, dreams could come true but *you* have to work at it and make it happen. When I first saw my sister-in-law working in the produce section at a local grocery store, I couldn't quite believe it. "*That's* what you do?" I asked her. "Yes", she answered, "and *you* have to do it too!"

Well, it's true; I did have to get a job in the same grocery store, as did most of my relatives from Pakistan. But, now that I have lived in Canada for five years, I have learned what it means to be a Canadian. Dreams do come true but there are no machines that make it happen. Since arriving in Canada, I now know about the importance of hard work, back there only one person of the family member (father) will work and all others will survive on it, but in here everyone have to work hard to fulfill their dreams, I learnt how it feels to my own house at the age of twenty, even though I work in a grocery store for minimum wage but still I am not dependent on others for my expenses , Being a Canadian I experienced the challenges of life from the age of 18 and overcome with them proudly. Canada has taught me how to be successful and face challenges in life. Canada has made me so mature. If I was still living back home in Pakistan, I would not be anywhere near as mature as I am now as I didn't value money, and had no understanding of how much work went into making it. Being a Canadian has made me appreciate the value of work. Canada *is* the land of dreams and comfort, but the secret behind is hard work which is hidden in the fancy name of moving to a western country. I am proud to be a Canadian because I am willing to work for it and make my dreams come true.

My comments:

1. Interesting.
2. Specific examples
3. Nice length
4. A good topic for the unit
5. However, she did not include ANY references to the texts. The story is good but she needs several references to what she has read.

Essay (Non-fiction): "The Canadian Personality" by Bruce Hutchison

"The Canadian Personality" by Bruce Hutchison was presented as a radio talk by the CBC, September 1, **1948**. Consider: Were these ideas true then? Are these ideas true today? In the talk show, Hutchison recognized the difficulty of defining the Canadian personality. He uses the voice of an American friend who is trying to capture the essence of being Canadian. **Below are only some of the paragraphs from the original radio talk show.** You will know where I skipped paragraphs because that is where I will have placed a …..

So, here are the parts of the essay that I want you to read….

Somewhere across this broad land of Canada tonight there is a lost and desperate man trying to find the smallest needle in the largest haystack in the world. He is one of the best American journalists in the business, he has covered important stories in countless countries, but his assignment in Canada has stumped him. His assignment is to discover, analyze, and spread on paper for the American public, the inner meaning of Canadian life.

Well, I did the best I could for the poor fellow. I talked to him all last night but when I had finished he was still pacing my room, aflame with the mystery of his mission and certain other stimulating refreshments I had provided – he was pacing the room at dawn and complaining that I had really told him nothing of Canada. "What I have to find," he cried out in his agony, "is the Canadian character, the Canadian personality, the Canadian dream."

When I last saw him, staggering into the sunrise, he hadn't found what he was looking for. And it suddenly occurred to me that I hadn't found it either, after half a century, that I probably wouldn't find it, that it may be forever undiscoverable. I am not surprised, therefore, when my American friend concludes that there actually is no Canadian character, personality, or dream.

Nevertheless, he was wrong. But he set me thinking. And the more I thought about this thing the more confused I became. Yet he was wrong.

….

Nothing of importance in life is definable. Once anything yields to definition you can be sure it isn't very important…..

We Canadians worry too much about our diversity. For it is an illusion ….to think that a nation grows strong by uniformity. Why, in the basic and most essential unit of mankind, in the family itself, diversity is the surest sign of strength and talent, the best guarantee of unity. No man in his senses would try to make his children all alike, and would mercifully extinguish them at birth if he thought they would resemble him when they grew up. What folly it is….to imagine

we have at last turned out a generation as uniform as a package of chewing gum and about as durable.

Nevertheless, as my bewildered American friend told me, it won't do to say that Canadians have strong and varied local characteristics in different parts of the country. That won't prove the existence of a national character. You must be able to prove that throughout the country there are certain dominant, widely shared and fully accepted characteristics, instincts, and deep feelings….. That is where the argument about our national character always collapses, as I have seen it collapse, over and over again, usually late at night amid a despairing clink of glasses, from Victoria to Halifax.

….

Despite everything, however, I think we can begin now to detect some of the special characteristics common to all Canadians, and add them up to something.

First, and most obvious, is our national humility. We are a people bounded on one side by the northern lights and on the other by an inferiority complex just as vivid, a people distracted by the mossy grandeur of the world from which we came and by the power, wealth, and fury of our American neighbors. We are the last people to realize, and the first to deny, the material achievements of the Canadian nation, which all the rest of the world has grasped and envied. Self-deprecation is our great national habit…. Never has there been a people in all history which has accomplished so much as the Canadian people and thought so little of it…..We write everything small if it's Canadian.

This, perhaps, lies close to the root of another national characteristic – we are a conservative and steady people, hardly daring to believe in our own capacity in the more complex affairs of the statecraft…. The Canadian audience at a political meeting is the most dead-panned ever known…and our politicians truly reflect us in their….positive terror of color and flair.

And we are a lonely people, isolated from one another, in a land where the largest city is a frail wink of lights in the darkness of night.

Lonely, and awed by the immensity of space around us, by the cold sweep of the prairies, by the stark presence of mountains leaning upon us, by the empty sea at our door, and by the fierce northern climate, which colors and toughens the weather of our spirit. And we are closer to the soil still, all of us, even in our cities, than the people of any other great industrial and urban nation.

We are more aware than others of the central physical fact of the earth, of growth, of harvest and decay. This land sense dominates all our national thinking, our politics, our

economic system, and our personal habits. It makes our artists instinctively rush out to paint, not the abstractions of other artists, but the hard material of rock and pine tree.

This deep instinct for the land, our constant feeling of struggle against a harsh nature – this and our concentration on the mere task of survival, must be one of the things that makes us an unimaginative people…..lacking in humor. (We haven't even developed the great Canadian joke yet or learned to laugh at ourselves.) It may turn out that we are really filled with fire, poetry, and laughter, which we have repressed, thinking it inferior to other peoples', and perhaps these things will erupt some day, with shattering violence…..

On the evidence so far you might almost say that we have constructed a national character by refusing to construct one. ….and perhaps the refusal to admit achievement is an achievement in itself.

But there is something about us more important and more distinctive than any of these obvious qualities.

We are among the few peoples still in the first throes of collective growth. While older peoples have settled down and accepted certain conventions, conditions, attitudes, and limitations as permanent, we accept nothing, least of all limitations. We live in a constant expectation of change, which we don't particularly relish and rather suspect, but cannot avoid. We have, every one of us, the feeling that we are involved in a process of perpetual expansion, development, and revision, whose end we cannot see.

We have the feeling, not of an old and settled resident, in his father's house, but of a young man building a new house for himself, without any clear plan in his head and wondering how large his future family will be…This, I think, is the universal and most distinctive feature of a Canadian. We are, above all, a building people, a nation of beavers.

But, my American friend says, all this does not add up to a national character, and hence he concludes there is none. All right, then. We have failed to define that character, as I told you we would. But consider this: We have built here against every obstacle of geography, economics, racial division, and the magnet of our American neighbor – we have built here the greatest nation of its population in all recorded history….

No political decision, no economic planning, will explain that. Something much more than politics or economics was at work – the unshakable will to make a nation, a home, a life of our own, for which no inconvenience was too great, from which no temptation could swerve us – a dim, impalpable, and dumb thing beyond our power to express or even name.

….

We quarrel about methods, political theories, economic systems, but such things do not make up a national character. Our character is not being built on them but on something much larger, a truly common denominator, the space, the beauty, and the free life of Canada itself.

Well, I wonder what haystack my American friend is searching in tonight for a needle which he could not recognize even if he found it.

Rubric for Response Journal Entries

Name: _____ Due Date: _____

Number of Responses to be Assessed (Drafts): _____

85-100% (17-20)	Journal responses are insightful and perceptive, connecting personal experience to the text and making observations and judgements. There is a thoughtful interpretation of what was read, heard, or viewed. Some entries go beyond personal experiences to connect to others and the world around you. Sometimes critical questions are asked. Effort is made to write about a page or more (300 words or more). Not every entry needs to be at this level for students to be rated at a 5 level.
75-85% (15-16.5)	Journal responses are mostly thoughtful. They show personal involvement with and understanding of the text, and make reasonable observations and judgements. They reflect an understanding of what is being viewed, read, or heard. Effort is made to write about a page (300 words).
65-75% (13-14.5)	Journal responses meet basic expectations but some responses are missing necessary detail or include unnecessary information. The writer relates to or identifies with what is being viewed, read, or heard, but only makes general comments about the text. Entries do not critically assess the ideas being presented. Entries are too short.
50-65% (10-12.5)	Journal responses are general and not expanded upon or may ramble repetitively without clear connections. The writer may empathize with or judge the text, but not consider the context or deeper meanings of the text. There are only vague references to the text itself. Entries are too short.
49% or lower (0-9.5)	Journal responses are incomplete, unclear, or show little effort or insight. The writer occasionally makes observations about the text but these are vague and unsupported. Entries are often simply summaries or retellings of the events in the text. The writer may rate the text, but give little or no support for assertions, judgements or personal opinion. Entries are too short.

Comments:

/20 + _____ (number of entries submitted) = /30

Consent Form

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN

Computer Systems Technology
Saskatchewan Polytechnic

Research Project: Reading comprehension and educational frameworks as a basis for predictions of student success

Investigators: Dr. Gordon McCalla, Professor, Department of Computer Science (966-4902), mccalla@cs.usask.ca
Terry Peckham, Department of Computer Science (966-2666), tep578@mail.usask.ca
Terry Peckham, Computer Systems Technology(659-4265), terry.peckham@saskpolytech.ca

This consent form should give you the basic idea of what the study is about and what your participation will involve. If you would like more detail about something mentioned here, or information is not included here, please contact one of the study investigators listed above. Please take the time to read this form carefully and to understand any accompanying information. Any involvement in this study is purely voluntary and is to be considered an optional activity. There are no known risks involved with this study.

For assignment 1 in this course, you will be accessing an online environment in which you read articles and answer questions about them. Your teacher will give you marks for each of your answers, which will be part of your grade for this module. This will happen regardless of whether or not you consent to take part in the study, since this is part of your workload in the course.

However, if you do consent to take part in the study, let us explain what it is about. The purpose of this study is to try to gain an understanding about how individual learners differ in how they use online material to answer questions. To this end, the experimenters will later be able to access a record of your online behaviour as you read and answered questions, and will also have access to your marks on each question. In addition, if you consent to take part, you will be asked to fill out a short questionnaire prior to beginning the study to gather background information about yourself and your prior knowledge, if any, of the subject matter that you will be reading and answering questions on. And, you will be asked to fill out an additional short questionnaire at the end. Further, after your participation in the study, you will be given more information about the purpose and goals of the study.

Rest assured if you do consent to take part in the study, that you will be completely anonymous to the experimenters and the wider world, and only known by a computer-generated identifier that cannot be traced back to you (see below). Any information collected in this study will be kept confidential and not shared with anyone outside of the research team. Any results that are published will mainly be in statistical form. In cases where a particular student's behaviour might be referred to in a publication, we will ensure that there will be no way any individual can be identified. As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled. This summary will outline the research and discuss our findings and recommendations.

A further reward for participation will be in the form of a lottery. This lottery will take all of the participants who have consented to take part in this experiment and draw a random winner from this pool. The prize will be awarded after the experiment has been completed and all of the data collected.

Your continued participation throughout the experiment should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact one of the investigators listed above.

If you do not participate in this study, the data automatically collected by the learning environment will be discarded and deleted upon your completion of the module.

By signing below, you indicate that you have understood to your satisfaction the information regarding participation in the research project and agree to participate. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Dr. Gordon McCalla, Professor, Dept. Computer Science (306) 966-4902
  mccalla@cs.usask.ca
  or
- Research Ethics Office University of Saskatchewan (306) 966-2084

Following your acceptance of this consent form you will be asked to create a user id and password in order to login to the study website. The system will email you a link to the study after you have created a user id and password. The passwords will be encrypted using a one way hash algorithm (MD5) so that no one will be able to use or figure out your password. When you have finished entering your user id and password a random computer generated id will be created and associated with all of your data so that your name and other information is not directly associated with you. If you should choose to withdraw from the study at any time, just enter your username and password into the website again and you can choose the option to withdraw at your convenience.

If you would like a copy of this consent form you may access it from our experiment website and print it through your browser, or contact one of the investigators listed above. This research

project has been approved on ethical grounds by the University of Saskatchewan Research Ethics Board. Any questions regarding your rights as a participant may be addressed to that committee through the Research Ethics Office ethics.office@usask.ca (306) 966-2975. Out of town participants may call toll free (888) 966-2975.

Print Name _____

Signature_____ Date: _____

Statistics Section

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	3.028	3	1.009	54.411	2.43E-11	88.87
within groups	0.482	26	0.019			11.13
total	3.51	29				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.13446	0.5182	0.5518	0.866154	-	-
Gabriel comparison interval	0.086	0.193	0.122	0.076	-	-
n	10	2	5	13	-	-

Statistics for Bloom Level 1

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-	0.2894	0.20464	0.15715
60,30,10	0.3837*	-	0.31259	0.28378
70,20,10	0.4173*	0.0336	-	0.19661
80,10,10	0.7317*	0.348*	0.31435*	-

Tukey-Kramer Bloom Level 1

	sum of squares	degrees of freedom	mean square	Fs	p	variance component (%)
among groups	0.628	2	0.314	51.904	2.00E-8	91.43
within groups	0.115	19	0.00605			8.57
total	0.743	21				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.335967	-	0.633767	0.821688	-	-
Gabriel comparison interval	0.083	-	0.083	0.036	-	-
n	3	-	3	16	-	-

Statistics Bloom Level 2

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0.16135	0.12433
60,30,10				
70,20,10	0.2978*		-	0.12433
80,10,10	0.4857*		0.18792*	-

Tukey-Kramer Bloom Level 2

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	2.407	2	1.203	110.714	3.38E-11	94.04
within groups	0.207	19	0.011			5.96
total	2.613	21				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.173711	-	0.568225	0.904622	-	-
Gabriel comparison interval	0.064	-	0.096	0.064	-	-
n	9	-	4	9	-	-

Statistics Bloom Level 6

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0.15917	0.12486
60,30,10				
70,20,10	0.3945*		-	0.15917
80,10,10	0.7309*		0.3364*	-

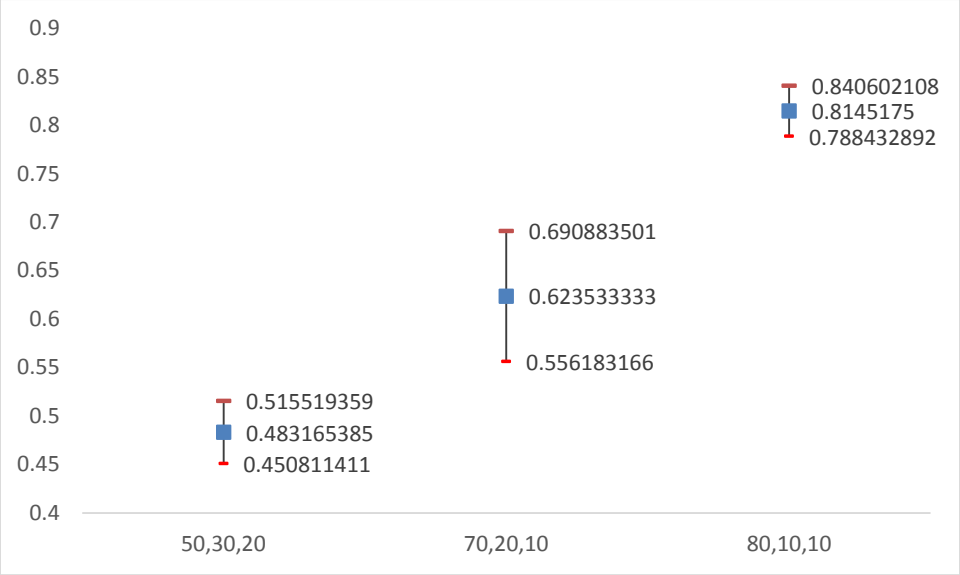
Tukey-Kramer Bloom Level 6

	sum of squares	**degrees of freedom**	**mean square**	**Fs**	**p**	**variance component (%)**
among groups	1.75	2	0.875	96.117	1.13E-20	82.67
within groups	0.628	69	0.009104			17.33
total	2.378	71				
	50,30,20	**60,30,10**	**70,20,10**	**80,10,10**		
mean	0.483165	-	0.623533	0.814518	-	-
Gabriel comparison interval	0.032	-	0.067	0.026	-	-
n	26	-	6	40	-	-

Statistics Keyboard Bloom Level 6

	50,30,20	60,30,10	70,20,10	80,10,10
50,30,20	-		0.10352	0.05758
60,30,10				
70,20,10	0.14037*		-	0.10007
80,10,10	0.3314*		0.19098*	-

Tukey-Kramer Keyboard Bloom Level 6



CGI for Keyboard Bloom Level 6

Appendix 4 Educational Taxonomies

Marzano's Taxonomy

There are several problems that exist within Bloom's Taxonomy. First, the taxonomy primarily focuses on educational outcomes found in the cognitive domain and tends to avoid the psychomotor and affective domains [106]. Second, the taxonomy does not sufficiently differentiate between the types of knowledge it presents and the processes by which each type of knowledge is retrieved or processed [106]. Last, since the taxonomy is organized as a hierarchy using the degree of difficulty as a marker, there is some controversy over its validity [106]. For example, there are cases where some knowledge types are assumed to be at a higher level within the hierarchy but involve the performance of tasks classified at lower levels in the hierarchy.

In response to shortfalls found within Bloom's taxonomy, Marzano and Kendall [107] in 2007 introduced their taxonomy of educational objectives that is similar yet different from Bloom's Taxonomy (see figure 3). Marzano's premise is that knowledge use is affected by three systems: the cognitive system, the metacognitive system and the self-system [107]. When an individual is faced with some new situation, the self-system must determine if it is better to continue with the current behaviour or to adapt some new behaviour. The metacognitive system then tries to set the goals that are needed to achieve the desired outcome and then monitor those goals. The cognitive system processes all the necessary information required to complete the task that is obtained from the knowledge system [107].

The lowest level of learning in Marzano's Taxonomy is to obtain knowledge. Knowledge is divided into three distinct types: information, mental procedures, and physical procedures. Information is defined as the "substance" that we think about and includes

generalizations, theories, data, vocabulary, etc. [107]. Mental procedures involve how we classify, analyze and apply the information. Physical procedures involve the skills we need to carry out tasks [107]. These include items such as typing, hand writing, sports, etc. to name a few.

How the information is processed is affected by the cognitive system. The cognitive system processes the information through one of the four cognitive stages: knowledge retrieval, comprehension, analysis, and knowledge utilization [107]. Like the knowledge component in Bloom's Taxonomy, knowledge retrieval involves obtaining information from memory. At this level, individuals are simply recalling facts, sequences or processes that they have memorized. Comprehension is a higher level of learning and involves properly identifying which information is important for the current task and making sure it is properly categorized. For example, a student learning about Napoleon's conquest of Russia should be required to remember which route was taken but not the numbers or types of weapons taken into battle. Since context is important, the type of information needed will vary by the task at hand. Analysis involves learners being able to take the knowledge they have and create new insights, inventing new ways of implementing knowledge in novel situations. Lastly, knowledge utilization involves the application of knowledge to carry out project-based learning since it involves skills required by people to accomplish a specific task. Project-based learning is a teaching method where students gain knowledge and skills by working for an extended period of time to investigate and respond to a complex question, problem or challenge [108].

Each of the four cognitive categories found in Marzano's Taxonomy also contains various sub-categories or skills that are found within the category. These 14 sub-categories, as seen in Figure A1, provide a richer set of categories for creating educational objectives and their

corresponding educational activities. More information about the sub-levels can be obtained from [107].

M,Cognitive System	4) Knowledge Utilization	4.Investigating
		3.Experimenting
		2.Problem-Solving
		1.Decision-Making
	3) Analysis	5.Specifying
		4.Generalizing
		3.Analyzing Errors
		2.Classifying
		1.Matching
	2) Comprehension	2.Symbolizing
		1.Integrating
	1) Retrieval	3.Executing
		2.Recalling
		1.Recognizing

Figure A4.1 Marzano's Cognitive Domain

Each of Bloom's categories for the cognitive domain can be mapped onto one of the categories for Marzano's cognitive domain. However, there is a one-to-many mapping between Bloom's and Marzano's taxonomy [106]. The method required to solve a problem within Bloom's taxonomy does not always equate to the same level that is defined within Marzano. So in practice we find that a problem categorized as Bloom level 2 (understanding) may equate to Marzano's level 2 (comprehension) or to Marzano's level 1 (knowledge) depending on the context of the problem. Marzano offers the same patterns of predictability as the Bloom level 1 example discussed in experiment 1. In experiment 2, both Bloom and Marzano seem to be no different from one another. However, not all Bloom level 1 questions map to Marzano level 1. For example, question 1 from experiment 1B mapped from Bloom level 2 to Marzano level 1 and question 5 from the same experiment mapped to Bloom level 3 to Marzano level 1. In

experiment 2A, question 2 mapped from Bloom level 1 to Marzano level 2. The fact that there is not a simple one-to-one mapping between Marzano and Bloom underscores the fact that there are differences in how the two taxonomies perform their labeling.

Marzano Taxonomy – Thinking Processes with Design Verbs

Level of Difficulty	Mental Process	Verbs and Phrases	
Level 4 Knowledge Utilization	Decision Making	Select the best among the following alternatives Which among the following would be the best…	What is the best way… Which of these is more suitable Decide
	Problem Solving	Solve How would you overcome… Adapt Develop a strategy to…	Figure out a way to… How will you reach your goal under these conditions…
	Experimenting	Experiment Generate and test Test the idea that What would happen if… How would you test that	How would you determine if… How can this be explained Based on the experiment, what can be predicted
	Investigating	Investigate Research Find out about Take a position on	What are the differing features of… How did this happen Why did this happen What would happen if
Level 3 Analysis	Matching/ Comparative Analysis	Categorize Compare & contrast Differentiate Discriminate	Distinguish Sort Create an analogy Create a metaphor
	Classifying	Classify Organize Sort	Identify a broader category Identify categories Identify different types
	Analyzing Errors	Identify errors Identify problems Identify issues Identify misunderstandings Assess	Critique Diagnose Evaluate Edit Revise
	Generalizing	Generalize What conclusions can be drawn What inferences can be made Create a generalization	Create a principle Create a rule Trace the development of… Form conclusions
	Specifying	Make and defend Predict Judge Deduce	What would have to happen Develop an argument for Under what conditions
Level 2 Comprehension	Integrating	Describe how or why Describe the key parts of Describe the effects Describe the relationship between	Explain ways in which Paraphrase Summarize
	Symbolizing	Symbolize Depict Represent Illustrate Draw	Show Use Models Diagram Chart
Level 1 Retrieval	Recognizing	Recognize (form a list) Select from (a list)	Identify (from a list) Determine (if the following statements are true)
	Recalling	Exemplify Name List Label State	Describe Identify who Describe what Identify when
	Executing	Use Demonstrate Show	Make Complete Draft

source: http://www.sweethomeschools.com/SDMaterials/TaxonomyVerbList.pdf

Source: https://icbg.files.wordpress.com/2012/06/4-action-verbs-marzano-ii.pdf

Blooms Taxonomy

Bloom's Taxonomy was originally developed to overcome several problems found within educational organizations. First, it provides a framework to facilitate the exchange of comparable test items between faculty, universities, etc. Second, Bloom viewed the taxonomy as a common language about the learning goals that facilitated communication across persons, subject matter, and grade levels. Third, it could also serve as the basis for categorizing a particular course or curriculum with respect to broad educational goals, such as those currently found in national, provincial and local standards. Lastly, it can provide a range of educational possibilities against which the limited breadth and depth of any particular educational course or curriculum could be contrasted [109] [8]. Since learning outcomes and learning steps are created at specific Bloom levels it becomes possible to compare and contrast courses and results.

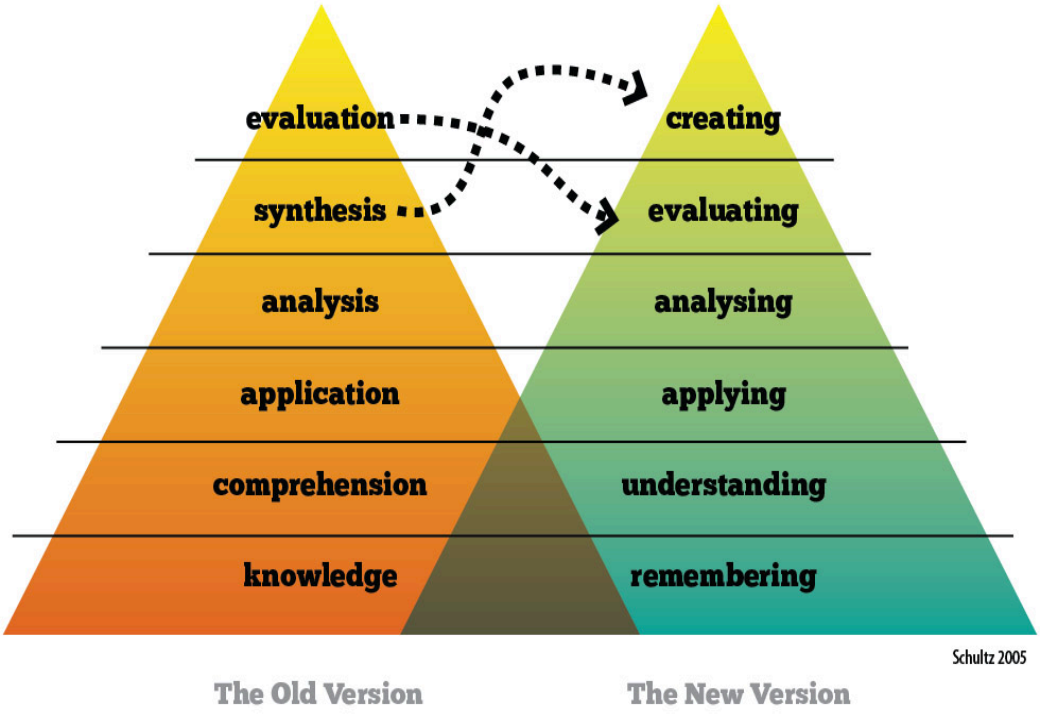
Bloom's taxonomy provides us with a framework where we can compare and contrast curriculum against the student interaction data logged by our system. Intuitively we can assume that for difficult problems as categorized by Bloom's taxonomy, a student must perform a deep methodical reading style so that they can create and synthesize connections that exist within the corpus of information being analyzed. But is this deep methodical reading style the best style for all levels of difficulty as defined within Bloom's taxonomy? We hypothesize that for tasks at different levels of Bloom's taxonomy, there will be effective and ineffective reading styles that emerge from our logged student interaction data.

Bloom's Taxonomy [9] and its subsequent revision [8] are comprised of three overlapping domains: cognitive, affective and psychomotor. The affective domain is comprised

of emotions, attitudes and values. The psychomotor domain is comprised of physical skill mastery, coordination, etc. The cognitive domain provides a method to classify educational objectives that relate to knowledge [7]. Within the cognitive domain are six hierarchical levels in order of increasing complexity: knowledge, comprehension, application, analysis, synthesis and evaluation (as revised by Anderson et. al. [8]). The first three levels are considered to be foundational learning and are based upon the ability to know and apply factual knowledge [7]. The last three levels are considered higher level learning that is more abstract in nature [110]. Bloom had originally assumed that you could not achieve the higher levels without first mastering the lower levels of the hierarchy [9]. However, it appears that it is possible to work at the higher levels on some topics without first mastering the lower levels [8].

Wankat and Oreovicz [111] provide some examples of how to apply Bloom's taxonomy to an engineering domain. Knowledge or recall involves the descriptions, definitions, generalizations and other routine information about a topic. Comprehension involves understanding the technical representations of a topic including the translation, interpretation and extrapolation of that topic. Application involves the use of topical abstractions in explicit situations such as the use of rules, procedures and theories to perform some computation. Analysis involves breaking a problem into its principal parts in order to highlight any prerequisite content hierarchy / properties. Furthermore, connections and structure found within the content are defined and clarified. Synthesis involves putting together all the constituent parts of a problem into a coherent system or solution. This can be very difficult since the process is open-ended and there may be many possible solutions to the problem. Lastly, evaluation can involve making conclusions about the value of materials used in a project or the methods used in that project. There is a need to satisfy specific criteria or use some standard of appraisal.

Through the use of the different levels of Bloom's Taxonomy and questions that are appropriately couched within the framework, it is possible to help learners to overcome various problems at each level of the taxonomy.



Source: http://www.psia-nw.org/newsletter-articles/blooms-taxonomy-levels-of-understanding/

Active verbs developed based on Bloom's Taxonomy

Knowledge	Understand	Apply	Analyze	Evaluate	Create
define	explain	solve	analyze	reframe	design
identify	describe	apply	compare	criticize	compose
describe	interpret	illustrate	classify	evaluate	create
label	paraphrase	modify	contrast	order	plan
list	summarize	use	distinguish	appraise	combine
name	classify	calculate	infer	judge	formulate
state	compare	change	separate	support	invent
match	differentiate	choose	explain	compare	hypothesize
recognize	discuss	demonstrate	select	decide	substitute
select	distinguish	discover	categorize	discriminate	write
examine	extend	experiment	connect	recommend	compile
locate	predict	relate	differentiate	summarize	construct
memorize	associate	show	discriminate	assess	develop
quote	contrast	sketch	divide	choose	generalize
recall	convert	complete	order	convince	integrate
reproduce	demonstrate	construct	point out	defend	modify
tabulate	estimate	dramatize	prioritize	estimate	organize
tell	express	interpret	subdivide	find errors	prepare
copy	Identify	Manipulate	survey	grade	produce
discover	indicate	Paint	advertise	measure	rearrange
duplicate	Infer	Prepare	appraise	predict	rewrite
enumerate	relate	produce	Break down	rank	role-play

Source:http://www.msjc.edu/CollegeInformation/Administration/Committees/Curriculum Committee/Documents/Blooms%20Taxonomy%2010%20-%2011.doc