ABSTRACT

AN AUTOMATED FEEDBACK SYSTEM TO SUPPORT STUDENT LEARNING OF CONCEPTUAL KNOWLEDGE IN WRITING-TO-LEARN ACTIVITIES

by
Ye Xiong

As a pedagogical strategy, Writing-to-Learn (WTL) intends to use writing to improve students' understanding of course content. However, most of the existing feedback systems for writing are mainly focused on improving students' writing skills rather than their conceptual development. In this dissertation, an automatic approach is proposed to generate timely, actionable, and individualized feedback based on comparing knowledge representations extracted from lecture slides and individual students' writing assignments. The novelty of the proposed approach lies in the feedback generation: to help students assimilate new knowledge into their existing knowledge better, their current knowledge is modeled as a set of matching concepts; suggested concepts and concept relationships for inclusion are generated as feedback by combining two factors, i.e., importance and relevance, of feedback candidates to the matching concepts in the domain knowledge. In the prototype system, a student can request feedback many times; each set of feedback is generated for a corresponding assignment draft to support their learning of conceptual knowledge during the iterative process of writing an assignment.

This research conducts a repeated measures study across two semesters (N=88) to understand how students perceive the proposed system, explore how students use the automated feedback, and investigate the effects of the automated feedback on student learning. Survey results show that the feedback is perceived as relevant (78.4%), easy to understand (82.9%), accurate (76.1%) and useful (79.5%); survey results also find that the

proposed system makes it easier to study course concepts (80.7%) and is useful in learning course concepts (77.3%). Based on the log analysis of students' actual usage of the system, all participants request feedback at least once when using the proposed system. After requesting feedback, 83 out of 88 participants revise their assignments. Analyses of students' submitted assignments reveal that more course concepts and concept relationships are included when completed using the proposed system. Collectively, these results show that the proposed automated feedback prototype system contributes to students incorporating more course concepts and concept relationships into their writing assignments, thus supports their learning of conceptual knowledge in a WTL activity.

AN AUTOMATED FEEDBACK SYSTEM TO SUPPORT STUDENT LEARNING OF CONCEPTUAL KNOWLEDGE IN WRITING-TO-LEARN ACTIVITIES

by
Ye Xiong

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Information Systems

Department of Informatics

August 2020

APPROVAL PAGE

AN AUTOMATED FEEDBACK SYSTEM TO SUPPORT STUDENT LEARNING OF CONCEPTUAL KNOWLEDGE IN WRITING-TO-LEARN ACTIVITIES

Ye Xiong

Dr. Yi-fang Brook Wu, Dissertation Advisor — Date
Associate Professor, Department of Informatics, NJIT

Dr. S. Roxanne Hiltz, Committee Member — Date
Distinguished Professor Emerita, Department of Informatics, NJIT

Dr. Michael Bieber, Committee Member — Date
Professor Emeritus, Department of Informatics, NJIT

Dr. Shaohua Wang, Committee Member — Date
Assistant Professor, Department of Informatics, NJIT

Dr. Bodong Chen, Committee Member — Date
Associate Professor, Department of Curriculum and Instruction,
University of Minnesota - Twin Cities

BIOGRAPHICAL SKETCH

Author: Ye Xiong

Degree: Doctor of Philosophy

Date: August 2020

Undergraduate and Graduate Education:

- Doctor of Philosophy in Information Systems,
  New Jersey Institute of Technology, Newark, NJ, 2020

- Master of Science in Business and Information Systems,
  New Jersey Institute of Technology, Newark, NJ, 2013

- Bachelor of Arts in English,
  Beijing Foreign Studies University, Beijing, P. R. China, 2009

Major: Information Systems

Presentations and Publications:

Xiong, Y. and Y.-f. B. Wu. (2019). Understanding Students' Perceptions of an Automated Feedback System: An Empirical Study Based on UTAUT. In *Proceedings of the 25th Americas Conference on Information Systems* (pp. 2221-2230). Cancún, Mexico.

Xiong, Y. and Y.-f. B. Wu. (2019). An Automated Feedback System to Support Student Learning in Writing-to-Learn Activities. In *Proceedings of the Sixth ACM Conference on Learning@ Scale* (pp. 1-4). Chicago, IL, USA.

Xiong, Y. and Y.-f. B. Wu. (2017). Write-and-Learn: Promoting Meaningful Learning through Concept Map-based Formative Feedback on Writing Assignments. In *Proceedings of the Seventh International Learning Analytics and Knowledge Conference* (pp. 552-553). Vancouver, BC, Canada.

Duan, L. and Y. Xiong. (2015). Big Data Analytics and Business Analytics. *Journal of Management Analytics* 2.1 (2015): 1-21.

Xiong, Y., T. Alelyani, R. Collins, and S. R. Hiltz. (2015). Not Just Silly Cat Videos: Exploring Student Knowledge Sharing via Social Media. In *Proceedings of the 21st Americas Conference on Information Systems* (pp. 3368-3381). Puerto Rico.

This dissertation is dedicated to my beloved family.

致我挚爱的家人

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my dissertation advisor, Dr. Yi-fang Brook Wu, for her constant help and guidance throughout my doctoral studies. No matter how busy she was, Dr. Wu has always found time to give me prompt and great advice. Her guidance and encouragement have kept me moving forward even when the end goal seemed unattainable. I would also like to express my sincerest gratitude to Dr. Starr Roxanne Hiltz for being a tireless and invaluable professor, who always made time to answer my questions and review my work. Special thanks are given to Dr. Michael Bieber for all his great help and advice since I enrolled into the program. My sincerest gratitude is also expressed to all my other committee members, Dr. Shaohua Wang and Dr. Bodong Chen, for their insightful feedback and comments on my dissertation research. This work would not have been possible without all their valuable advice and support.

I would like to express my heartfelt appreciation to Dr. Richard Egan for his extensive support with my data collection. Also, I am deeply grateful to Dr. Lian Duan, Dr. Katia Passerini, Dr. Michael Chumer, Dr. Jerry L. Fjermestad, Dr. Vassilka Kirova, Dr. Sameh A. Sabet, Dr. Michele Scott, and Mr. Sathish Rajamani who offer me all kinds of help and support during my graduate studies. Special thanks are given to our department secretary, Ms. Patricia B. Lundberg, for all her great assistance.

Finally, I would also like to express my sincerest gratitude to Dr. Yang Liu, who provides me great technical support whenever I need it and deserves more recognition than a mention on this page. I wish to sincerely thank Dr. Regina Collins, Dr. Mingzhu Zhu, Dr. Christopher Markson, Dr. Yanchi Liu, and many other fellow students for their help and encouragement throughout this journey.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

The nation's higher education system is being challenged by a global knowledge economy that requires a literate workforce (Darling-Hammond, 2015). Students need to acquire knowledge of basic concepts of a discipline to build up a solid conceptual foundation for successful academic and professional development. Current efforts to reform education are largely targeted at increasing students' subject matter learning, especially knowledge and understanding of course concepts. As a pedagogical strategy, Writing-to-Learn (WTL) (Zinsser, 1989) has been widely adopted to improve students' deep understanding of conceptual knowledge. In the context of WTL, writing is used to help students reflect and develop their understanding of course content and concepts (Reynolds et al., 2012). Through different WTL assignments (e.g., summary, reflective essay), students can think through core concepts or ideas presented in a course and reflect upon what they know about course topics.

Feedback is one of the most significant interventions in learning (Hattie & Timperley, 2007), and feedback on writing is widely acknowledged to offer considerable learning benefits. Most existing feedback systems for writing (Foltz & Rosenstein, 2015; Villalon et al., 2008) are mainly focused on improving students' writing skills rather than their conceptual development. Without *timely* and individualized formative feedback, students lack much-needed information to improve their conceptual understanding *during* WTL activities. On the other hand, it is too labor intensive for an instructor to provide timely and individualized formative feedback throughout the duration of an assignment for an entire class of students.

The motivation of this research is to explore an automatic approach to provide students with formative feedback to promote their meaningful learning of conceptual knowledge in WTL activities. To make it a reality, it is important to understand how students learn meaningfully in general. According to the National Research Council report *How People Learn: Brain, Mind, Experience and School* (Bransford et al., 1999), people construct new knowledge and understanding based on what they already know and believe. New concept meanings are integrated into the cognitive structure to a greater or lesser extent, depending on how much effort the students make to seek this integration, and on the quantity and quality of their existing and relevant cognitive structures (Novak, 2002).

Meaningful learning takes place by the assimilation of new concepts and propositions into existing cognitive structures held by learners (Novak & Canas, 2007). Many cognitive science researchers consider the goal of meaningful learning to be the continued and organizational development of conceptual understanding to move learners from a novice state toward that of expertise (Romance & Vitale, 1999). In order to provide automated feedback, there is a need to represent what students know and then show the differences between domain or expert knowledge and a current student's knowledge state.

Previous studies show that concept maps, as a cognitive visualization and pedagogical tool to visualize the relationships among different concepts (Villalon & Calvo, 2011), are suitable to represent students' learning progress. Concept maps are regarded as a direct method of looking at the organization and structure of an individual's knowledge within a particular domain and at the fluency and efficiency with which the knowledge can be used (Williams, 1998). It is suggested that concept maps do capture a representative sample of conceptual knowledge and can differentiate well among fairly disparate levels of understanding based on the general homogeneity of the expert maps and their distinct variance from the student maps (Williams, 1998).

This research synthesizes theory and technology to explore if and how concept maps can be utilized one step further: from visualizing student learning progress to generating automated, timely, actionable, and individualized formative feedback on the development of conceptual knowledge in WTL activities. Lecture slides of a course subject are used to represent the domain-specific knowledge, because they essentially reflect all the major concepts that instructors intend to teach or students are expected to master for a course subject as suggested in related work (Atapattu et al., 2012; Gantayat & Iyer, 2011; Ono et al., 2011). Moreover, writing assignments in the context of WTL can reflect individual students' understanding of course content and concepts. Extended written responses are often regarded as an excellent means of determining how well students have understood certain concepts and can express their interrelationships (National Research Council, 2001).

Once the domain knowledge and student knowledge are represented, automated feedback can be generated based on the comparison of these two knowledge representations. In this research, the focus of the proposed system is to provide students with individualized formative feedback that suggests *what* concepts and concept relationships might be considered for addition to the assignment, not *how* they should be added. The prototype system cannot interact with a student about their writing like a human tutor, eliciting an appropriate response through carefully calibrated questions. However, it can make students aware of where the gaps lie in conceptual understanding and provide suggestions for the improvement of their coursework. In this student-centered learning environment, students can have the opportunities to learn meaningfully because of the purposeful integration of new knowledge with existing understanding.

1.2 Research Overview

The overarching objective of this research is to support student learning of conceptual knowledge in WTL activities by providing individual students with automated formative feedback. To achieve this goal, our research aims to explore: (1) how effective is the automated formative feedback generated with the help of concept maps constructed from lecture slides and writing assignments; (2) whether and how students utilize the automated formative feedback in WTL activities; (3) to what extent the automated formative feedback on writing assignments can affect student learning outcomes; and (4) do students recommend future use of the proposed automated feedback system? Why or why not?

This research begins with exploring how to generate automated feedback through the comparison between the two sets of concepts and concept relationships: one from the instructors' lecture slides and another from individual students' writing assignments. Concept maps in this research are used only as the tool to identify the key concepts and the concept relationships, and then derive automated feedback. Although the feasibility of comparing the similarity of two graphs is supported by previous studies (Andrews et al., 2009; De Souza et al., 2008), for the purpose of generating an itemized feedback list, there is no need to compute graph similarity of two concept maps in the proposed system; by using the map input, which includes the set of domain concepts and domain concept relationships, as well as the set of student concepts and student concept relationships, feedback can be automatically generated without the actual 'map,' i.e., the graphical representation. Thus, the graphic representation of concept maps is not implemented in the system for our research purposes and is beyond the scope of the study.

In the proposed system, the lecture slides used to extract domain concepts and domain concept relationships need to be well structured, which means that there is a title (or headline) on the top of each slide and text content related to the title in each slide. Also, the writing assignments used to extract student concepts and student concept relationships are typically designed to allow students to apply conceptual knowledge presented in a course. Only text content in the lecture slides and writing assignments are considered for our research purposes, which require that the lecture slides and writing assignments are mainly written in natural text and in the English language.

Once the domain knowledge representation and student knowledge representation are constructed, the task is to compare these two knowledge representations via the following two steps: (1) identifying matching concepts, i.e., identifying concepts appearing in both the student concept set and also the domain concept set; and (2) identifying matching concept relationships, i.e., identifying those concept relationships appearing in both the student concept relationship set and also in the domain concept relationship set with identical connecting concepts. Thus, automated feedback can be derived by comparing the two sets of concepts and concept relationships extracted from instructors' lecture slides and students' writing assignments, either finished or in-progress.

The novelty of our approach lies in the feedback generation, which is a design based on Ausubel's learning theory (Ausubel, 2000; Ausubel et al., 1968). To help students assimilate new knowledge into their existing knowledge, student knowledge is modeled as a set of matching concepts, and then suggested concepts and concept relationships are generated as feedback by combining two factors: importance and relevance of candidate concepts and concept relationships to the matching concepts in the domain knowledge representation. The resultant feedback is in the form of suggested concepts and concept relationships for inclusion for each input assignment draft. Thus, we consider the resultant automated feedback actionable and individualized.

The major contributions of this research include exploring an automatic approach to generate actionable and individualized formative feedback based on knowledge representations extracted from instructors' lecture slides and individual students' writing assignments, developing an automated feedback prototype system for a WTL environment that aims to support student learning of conceptual knowledge in the course of writing an assignment, and providing empirical evidence on how automated formative feedback can be provided to promote student learning in WTL activities. To achieve our research goals, the following research questions are explored:

RQ1. *How effective is the automated formative feedback generated with the help of concept maps constructed from lecture slides and writing assignments?*

RQ2. *Whether and how do students utilize the concept map-based formative feedback in WTL activities?*

RQ3. *To what extent can the concept map-based formative feedback on writing assignments affect student learning outcomes*?

RQ4. *Do students recommend future use of the proposed automated feedback system? Why or why not?*

1.3 Dissertation Outline

This dissertation is organized as follows. Chapter 1 introduces the background and motivation of the dissertation, and also presents an overview of this research. To provide a foundation in which to explore the educational value of the research, Chapter 2 presents a literature review of relevant learning theories and assessment principles, and then discusses related work of this research. Based on the open challenges identified, Chapter 3 presents the Write-and-Learn system architecture, and illustrates the interface of the prototype system, with a focus on the introduction of the formative feedback generation process. Chapter 4 discusses the results of preliminary studies including the evaluation of extracted domain concepts, think-aloud protocol analysis, and pilot study. Chapter 5 presents the findings of the repeated measures user study over the course of two semesters by providing descriptive statistics and quantitative data analysis. Finally, Chapter 6 summarizes the key research findings of the dissertation, discusses limitations and contributions of this research presented, and suggests avenues for future work.

CHAPTER 2

LITERATURE REVIEW

Learning science and technology combined with assessment theory lays a foundation for new and better ways to provide students with automated formative feedback in WTL activities. This research is grounded in cognitive sciences, supported by learning theories, and guided by assessment principles and practices. This chapter begins by introducing the pedagogical strategy of Writing-to-Learn, and then provides a review of the literature related to this study. Finally, the limitations of existing approaches are discussed, and our main research topics are identified.

2.1 Writing-To-Learn

Educational research suggests that writing is a task where higher cognitive functions, such as analysis and synthesis, can be fully developed (Emig, 1977). The act of writing has been recognized as high-impact learning strategy across disciplines (Kuh, 2008). Writing has been proven to be effective in promoting student learning, engagement, and success in relatively large enrollment face-to-face courses (Reynolds et al., 2012). The process of writing is important not only for learning about something or acquiring knowledge, but for constructing knowledge (Rivard, 1994). Prior research (Elbow, 1994) points out that it is helpful to distinguish between two different goals for writing: (1) *Learning-to-Write*, in which writing is the normal and conventional goal to demonstrate learning; and (2) *Writing-to-Learn* (also referred to as writing for learning), which is another important kind of writing and particularly effective at promoting learning and involvement in course materials. WTL is defined as a pedagogy that actively involves students across different disciplines in the construction of their own knowledge through writing (Carter, 2007; Comer et al., 2014; Deane Sorcinelli & Elbow, 1997).

As a pedagogical strategy, WTL has been extensively adopted to enhance knowledge acquisition and cognitive skill development in different disciplines (Reynolds et al., 2012). According to (Forsman, 1985), WTL is learning to think on paper which illustrates what a student already knows, and how his or her prior knowledge fits with new information being studied in a curriculum. Through various WTL assignments, students can think through core concepts or ideas presented in a course (Forsman, 1985). As summarized in (Reynolds et al., 2012), the types of WTL assignments include summary, reflective essay, synthesis, term paper, short paper, in-class writing, laboratory report, peer review, etc. These writing assignments represent a substantial component of undergraduate and graduate education (Morton, 2007), which can reflect what students know about course topics and develop higher-level cognitive processes that facilitate meaningful learning. It is argued that the importance of WTL should be highlighted as a significant pedagogical practice and encouraged more in massive open online courses (MOOCs) across disciplines (Comer et al., 2014).

2.2 Formative Feedback

In education, assessments are often used to obtain information about student learning and achievement. Formative assessment and summative assessment are two widely accepted approaches of assessment. This overview discusses the differences between them and the general principles for providing automated formative feedback, and finally introduces the existing systems or tools that provides feedback related to this research.

2.2.1 Formative Assessment

Different from summative assessment (i.e., assessment ***of*** learning) that is concerned with summarizing students' achievement status, formative assessment (i.e., assessment ***for*** learning), is regarded as a process, rather than a test, to continuously monitor, provide feedback, and respond to students' learning progress (Harlen & James, 1997). According to (Sadler, 1989, 1998), formative assessment is concerned with how judgments about the quality of student responses (performance, pieces, or works) can be used to shape their competence, and is specifically intended to generate feedback on performance to improve and accelerate learning.

As a key element in formative assessment, formative feedback is defined as information communicated to learners that is intended to modify their thinking or behavior for the purpose of improving learning (Hattie & Timperley, 2007; Shute, 2008). Formative feedback can take many forms (e.g., hints, error-flagging, correct response, and worked examples) (Shute, 2008), depending on different learning and instructional goals, research purposes, and methodological approaches (Ifenthaler, 2010). In general, formative feedback can be classified into two main functions: (1) directive feedback (also known as corrective feedback) that tells students what needs to be fixed or revised, such as right or wrong, overall percentage correct, try-again, and error-flagging, etc.; and (2) facilitative feedback (also known as elaborative feedback) that provides suggestions to guide students in their own revision and conceptualization, such as hints, cues, and prompts, etc. (Black & Wiliam, 1998).

Providing students with formative feedback has proven to be an effective strategy that is beneficial for student learning and crucial to improving knowledge and skill acquisition (Hattie & Timperley, 2007; Shute, 2008). First, formative feedback can signal a gap between a current level of performance and some desired level of performance throughout the learning process in the context of specific learning activities (Shute, 2008). Encouraging students to reflect on their work while they are engaging with the topic and task should have the most impact on students' understanding (Whitelock et al., 2015). Resolving the gap can also motivate higher levels of learning efforts (Shute, 2008). Students can use this much-needed information to determine which knowledge they need to study further and what adjustments in their thinking they need to make. Moreover, formative feedback can effectively reduce the cognitive load of learners, especially novice, struggling, and low-performing students (Shute, 2008). Results from prior work (Barnes & Stamper, 2010) suggest that students particularly need hints when they get stuck.

2.2.2 Computer-Based Feedback for Writing

The abundance of widely available computer-related technologies has exerted a significant impact on educational assessment. There is an increasing use of technology to support assignment delivery or submission and the medium for offering feedback (Whitelock, 2018). For example, the adoption of learning management systems (LMSs) or course management system (CMSs) can address high workload of assignment submission and grading (e.g., Moodle, Blackboard, WebCT, Canvas), as well as plagiarism detection (e.g., Turnitin) (Özbek, 2016), etc. In these systems, instructors can manually analyze these student writings or responses, assess their performance, and provide feedback or comments online after students submit their assignments or responses.

With the emergence of computer-based educational technologies, data mining and Natural Language Processing (NLP) techniques have been applied in the domain of education, especially automatic educational assessment. One alternative to feedback by instructors could be to implement computer-based feedback (Lachner et al., 2017). Drawn on multidisciplinary insights from computer science, linguistics, and educational data mining, a variety of computer-based systems or tools have been developed to automate the process of both scoring and providing feedback to serve the needs of different writing contexts. As suggested in (Madnani & Cahill, 2018), both scores and feedback in these systems are usually based on linguistic characteristics of the student discourse including but not limited to: (1) lower-level errors in response (such as grammatical or spelling errors in written responses); (2) discourse structure and organization of a piece of writing; and (3) relevance of the discourse to the question asked.

Most existing automated essay scoring (AES) systems, such as Holt Online Scoring, are focused on providing automated scoring rather than generating feedback (Nathawitharana et al., 2017). It is argued that these AES systems are mainly used to overcome time, cost and reliability issues in writing assessment (Dikli, 2006), and most of them offer little or no formative feedback to students other than the scores (Villalon et al., 2008). Also, there are automated essay evaluation (AEE) or automated writing evaluation (AWE) systems (Warschauer & Ware, 2006) in existence and some are commercially available, such as Criterion, MY Access, WriteToLearn, Summary Street, LightSIDE, OpenEssayist, BETSY, and WriteLab, etc. Although most AEE or AWE systems can provide students with feedback and assist instructors in expediting the feedback process, the primary concern of these systems is to evaluate writing proficiency and language skills, support improvements in writing motivation and writing quality (Palermo & Thomson, 2019; Wilson & Andrada, 2016).

In these systems, formative feedback provided is mainly in the form of descriptive information about a particular set of surface features of student writing (e.g., grammar, mechanics, style, structure, and coherence), with a focus on the assessment and improvement of writing skills rather than the development of conceptual knowledge for a specific course subject. For instance, Pearson's WriteToLearn system (Foltz & Rosenstein, 2015) provides students with writing exercises and automated feedback in terms of organization, word choice, and sentence fluency. Criterion, developed by Educational Testing Services (ETS), can generate a score for an essay and provide feedback on grammar, mechanics, style and organization (Attali & Burstein, 2004). Such feedback is focused on how to fix the incorrect and poor attributes of writings and how to improve writing skills by pointing out failings in the writing features. In (Kintsch et al., 2000), the system, State the Essence, was developed to help students learn how to write good summaries, where the feedback provided goes beyond other forms of automatic feedback, such as spelling and grammar check, and contains information including overall score, word length, adequate content coverage, as well as missing information. In their another attempt, the system Summary Street can ask students to write a summary and then provide feedback in terms of content, length, copying, spelling, redundancy, and irrelevancy (Kintsch et al., 2000).

2.2.3 Automated Formative Feedback for Writing-to-Learn

Providing students with feedback during the process of writing is crucial to the learning process (Villalon et al., 2008). Based on previous research, providing students with automated formative feedback on the development of conceptual knowledge in a WTL activity should take the following aspects into consideration.

Foremost, writing represents a unique mode of learning (Emig, 1977) and the purpose for writing activities should be taken into account in the development of feedback systems for writing (Villalon et al., 2008). In the context of WTL that intends to use writing to improve students' understanding of disciplinary content, writing is not only a tool for demonstrating learning, but also a tool for learning content, which can help students discover what they know about a topic and develop further understanding of a topic (Chatel, 1997). In WTL activities, greater emphasis should be given to students' subject matter learning where writing requires deep cognitive engagement with disciplinary concepts. As a result, the focus of formative feedback on WTL assignments is to promote the development of conceptual knowledge in a course subject.

Furthermore, according to (Nicol & Macfarlane-Dick, 2006), feedback needs to provide information about how a student's present state of learning or performance relates to the goals and standards. Several feedback systems for writing intend to engage users on matters of content. For instance, OpenEssayist developed by (Whitelock et al., 2015) processes open-text essays and offers feedback through key phrase extraction and extractive summarization. However, without a reference model of these goals and standards, such feedback contains no specific and actionable information on how to close the knowledge gap between student actual performance and faculty desired expectation. In (Kintsch et al., 2000), it is concluded that students are often dismayed at the multiplicity of problems to deal with, and many students need extensive and quite explicit guidance on how to make meaningful changes in revising their summaries.

Moreover, formative feedback can be provided in the form of facilitative feedback that shows suggestions to guide students in their own revision and conceptualization (Black & Wiliam, 1998) as mentioned earlier. Encouraging students to reflect on their work while they are engaging with the topic and task should have the most impact on students' understanding (Whitelock et al., 2015). Formative feedback needs to provide specific and actionable information so as to help individual students compare their own conceptual understanding with the intended learning outcomes, reflect on how they are conceptualizing the course content, rethink the reasons behind differences and similarities between their conceptualization and that from the reference model, find new concepts and concept relationships, and then use those reflections to construct their knowledge bases (Berlanga et al., 2012).

In addition, formative feedback needs to be provided based on the comparison between a student's course work and reference models that represent an expert's best-practice solution to complete the task, with respect to the expected learning outcomes (Ifenthaler, 2010). It is claimed that the predefined reference model can be built out of intended learning outcomes described in course materials, tutor notes, and curriculums, etc. (Berlanga et al., 2012). Students need to be given feedback that supports them in understanding task requirements, but also motivates them to believe they can make improvements on their coursework (Whitelock, 2018).

Finally, timely (often real-time) formative feedback is a fundamental component for supporting and regulating learning processes (Ifenthaler, 2010). It is pointed out that receiving quick and targeted feedback during the writing process is a critical support for learning from writing assignments (Ferster et al., 2012). In most cases, students have little opportunity to use directly the feedback they receive to close the performance gap, and greater emphasis might need to be given to providing feedback on work-in-progress (Nicol & Macfarlane Dick, 2006). But it is impossible for an instructor to do for an entire class of students at the time when it has the most effect (i.e., when students are engaging in or have just finished engaging in such activities) (Shute, 2008), and too labor intensive to do regularly in high school and introductory-level college classes which usually have a large number of students. Thus, the reference models should be generated automatically (or semi-automatically but at least with minimal human intervention), to allow students to reflect on their own work and understand their position in a knowledge domain (Berlanga et al., 2012). The following section discusses further why concept maps can be integrated into WTL activities and explores how formative feedback can be generated with the help of concept maps.

2.3 Concept Maps

This section includes the origin and definition of concept maps, the application of concept maps, and the approaches to construct concept maps from textual resources in the educational domain.

2.3.1 Definition of Concept Maps

Concept maps, introduced by Novak and Gowin (1984), represent a person's understanding of a topic by mapping concepts and their relationships in a graphical way (Villalon & Calvo, 2011). A concept map is composed of *nodes* that represent concepts, usually signified by nouns or noun phrases, and *links* connecting to nodes that represent the relationships between concepts. Each *node-link-node* triplet forms a *proposition* (Novak &

Cañas, 2008), which refers to a basic unit of statement about the object or event. "Concept" is defined by Novak and Gowin (1984) as a perceived regularity or pattern in events (i.e., "happenings") or objects (i.e., "things"), or records of events or objects. Villalon and Calvo (2010) point out that this definition is closely related to what is known in linguistics as "nouns" and validate the closely related definitions of nouns or noun phrases and concepts in their study. Consistent with the concept map research (Villalon & Calvo, 2011; Villalon et al., 2010), concepts in our research refer to nouns, compound nouns, or noun phrases covered in instructors' lecture slides or students' writing assignments for a course subject.

2.3.2 Application of Concept Maps

Concept maps have been regarded as one of the most common ways of representing cognitive structures. Research evidence has demonstrated the appropriateness of concept maps in eliciting knowledge (Cañas & Carvalho, 2004), organizing and representing knowledge within a domain (Croasdell et al., 2003), as well as evaluating learners' conceptual understanding (Jonassen et al., 1997). Over the past decade, there is a remarkable growth in the use of concept maps in education to assess and facilitate student conceptual understanding (Novak, 2005, 2010; Sengul & Senay, 2014).

As a tool to represent the structure of knowledge, concept maps have been widely applied in educational settings for different purposes. Foremost, concept maps can be used in assessment of learning, especially students' understanding of a domain's conceptual structure by using concept mapping tasks (Gouli et al., 2004; Lukasenko et al., 2010; Trumpower & Sarwar, 2010). In these tasks, students can arrange or label nodes and links to show relationships among multiple concepts in a domain (National Research Council, 2001). Moreover, concept maps can be used to facilitate meaningful learning and trigger reflection, especially in writing and reading activities, by making abstract knowledge and understanding visible in many different ways (Enger, 1996; Hay et al., 2008; O'donnell et al., 2002).

As for the integration of concept maps into writing, concept maps have been applied to structuring and organizing ideas in student writing in two ways First, concept maps can be used as a planning and organizing tool to facilitate writing when they are constructed prior to the start of writing. In such a scenario, concept maps can help students generate ideas, relate the ideas or content to each other, and visualize what is going to be written, as well as sequence the flow of writing. For instance, in (Al-Shaer, 2014), concept mapping is employed as pre-writing strategy to help learners better generate argumentative compositions. Second, concept maps can be used as a reflection tool after completion of writing. For example, in (Wan Mohamed & Omar, 2008), students are asked to construct concept maps based on their term papers after completing the papers and then do a reflection paper on how they felt using concept maps.

Existing work on the integration of concept maps into writing is primarily focused on the facilitation and reflection of writing, rather than the improvement of subject learning. For instance, in (Villalon & Calvo, 2009), researchers introduced a new approach to automatically extract a representation of the semantic information contained in student essays, so as to surface students' understanding about a topic. With the ways concept maps are used in these studies, either as a planning tool before writing or as a reflection tool after writing, little feedback is given in the course of writing. If formative feedback can provide actionable and individualized information on the development of conceptual knowledge and be automatically generated during WTL activities, it is useful for individual students to improve their conceptual understanding of a course subject.

2.3.3 Concept Mapping

Concept mapping (also called concept map mining) refers to the automatic or semi-automatic creation of concept maps from documents (Villalon & Calvo, 2008). As a constructivism-based learning strategy, concept mapping is a cognitive technique (Villalon & Calvo, 2011) to capture and examine human concepts, as well as to visualize the relationships or connections among different concepts. Concept mapping, as a means to make learning visible, can be utilized to embed the research on student learning (Hay et al., 2008).

First, A significant area of educational research is to gain a better understanding of how people learn (Hay, 2007). Based on Ausubel's learning theory (Ausubel, 2000; Ausubel et al., 1968), there is a distinction between no learning, rote learning, and meaningful learning: (1) no learning is indicated by an unchanged knowledge structure; (2) rote learning is indicated by some new or rejected concepts but no new links between prior and new knowledge; and (3) meaningful learning is indicated by significant revision to the knowledge structure (Ferrara & Butcher, 2011). In meaningful learning, students' prior and existing knowledge is more actively utilized to assimilate and make sense out of the new knowledge they are learning. During the process, the conceptual relations between new knowledge and existing knowledge are actively constructed (Glynn & Muth, 1994). Researchers have been exploring how to help students become active learners who seek to understand complex subject matter, and be better prepared to transfer what they have learned to new problems and settings (Bransford et al., 1999). Thanks to the Internet, knowledge is now all around us, which makes the memorization of facts or topics no longer necessary. Instead, the ability to use, manipulate, and apply that knowledge can differentiate high-performing students from the rest (Atkinson & Mayo, 2010).

Second, it is believed that the internal representation of knowledge resembles webs or networks of ideas that are organized and structured; and the more connections that exist among facts, ideas, and procedures, the better the understanding (Hiebert & Carpenter, 1992; Pintrich et al., 1993; Williams, 1998). As suggested in (National Research Council, 2001), one of the most important differences between experts and novices lies in how their knowledge is organized. Several studies on the examination of differences between experts and novices (Anderson, 1993; Ifenthaler, 2010) suggest that the conceptual organization of knowledge is the major characteristic of expert proficiency. Generally, experts organize their knowledge around core ideas or concepts and can see patterns of meaningful information that are not available to novices (Chi et al., 1981). Many cognitive science researchers consider the goal of meaningful learning to be the continued and organizational development of conceptual understanding to move learners from a novice state toward that of expertise (Romance & Vitale, 1999). Students who can categorize their knowledge and construct relationships between concepts are likely to promote expert-like thinking about a domain. A highly integrated knowledge structure signals the transition from novice to expert performance (Royer et al., 1993).

A considerable body of literature has demonstrated the feasibility of automatic or semi-automatic construction of concept maps and similar representations (e.g., topic maps, knowledge maps), from structured and unstructured textual data sources, including lecture slides (Atapattu et al., 2012, 2014a), textbooks (Larranaga et al., 2014; Olney et al., 2011), student essays (Villalon & Calvo, 2009, 2011), academic articles (Chen et al., 2008), domain handbooks (Hsieh et al., 2011), Wikipedia entries (Hartmann et al., 2012), etc. Some researchers follow the strict definition of a hierarchical concept map, while others use knowledge representations in more variable forms (Zubrinic et al., 2012), such as semantic networks, topic maps, knowledge maps, and mind maps, which have been utilized in knowledge management systems and are useful for knowledge codification, navigation, search and retrieval (Wang et al., 2011).

2.4 Summary

This research integrates insights from the fields of cognition and computation of language, learning analytics, and education theory. This chapter presented a review of prior literature related to this research from two major aspects. Foremost, to understand the theoretical foundations of this research, learning theories, assessment principles and related work were discussed in this chapter. Although a few systems can provide feedback for writing, most of them are focused on the assessment and improvement of students' writing skills. A review of existing feedback systems for writing was presented to support the underlying motivation for this dissertation, which is to improve student learning of conceptual knowledge in WTL activities. Moreover, to explore an automatic approach to generate meaningful feedback during the process of writing an assignment, the state-of-art concept mapping research was presented. With NLP techniques, it is feasible to provide students with automated feedback by the comparison between two knowledge representations: one from instructors' course materials such as lecture slides and another from students' writing assignments. The chapter concluded by discussing the gaps in knowledge and explaining the significance and feasibility of developing a new automated feedback system for a WTL environment. The next chapter illustrates the system architecture and interface of the proposed automated feedback system, and discusses the system implementation, with a focus on the approach to generate automated formative feedback to facilitate student learning in WTL activities.

CHAPTER 3

SYSTEM DESIGN AND IMPLEMENTATION

This research aims to develop an automated feedback system to provide individual students with automated feedback to assist their learning of conceptual knowledge in WTL activities. For this purpose, the first and critical step is to explore if and how formative feedback can be automatically generated based on the comparison of domain knowledge representation and student knowledge representation. The Write-and-Learn system is proposed to provide an environment in which students can prepare multiple drafts of assignments with feedback on the development of conceptual knowledge. This chapter presents an overview of the system architecture, with a focus on the methodology for generating automated feedback, discusses the system implementation, and illustrates how students interact with the prototype system in detail.

3.1 System Architecture

In this research, the proposed prototype system is developed as a free and open-source pedagogical tool that provides unique capabilities: generating automated, real-time, actionable and individualized formative feedback on students' writing assignments, so as to promote their learning of conceptual knowledge in a course subject. As illustrated in Figure 3.1, the system is composed of three major modules: (1) domain knowledge representation construction: a module which extracts domain concepts and domain concept relationships from instructors' lecture slides; (2) student knowledge representation construction: a module which extracts student concepts and student concept relationships from students' writing assignments; and (3) formative feedback generation: a module which provides automated feedback based on the comparison of two knowledge representations.

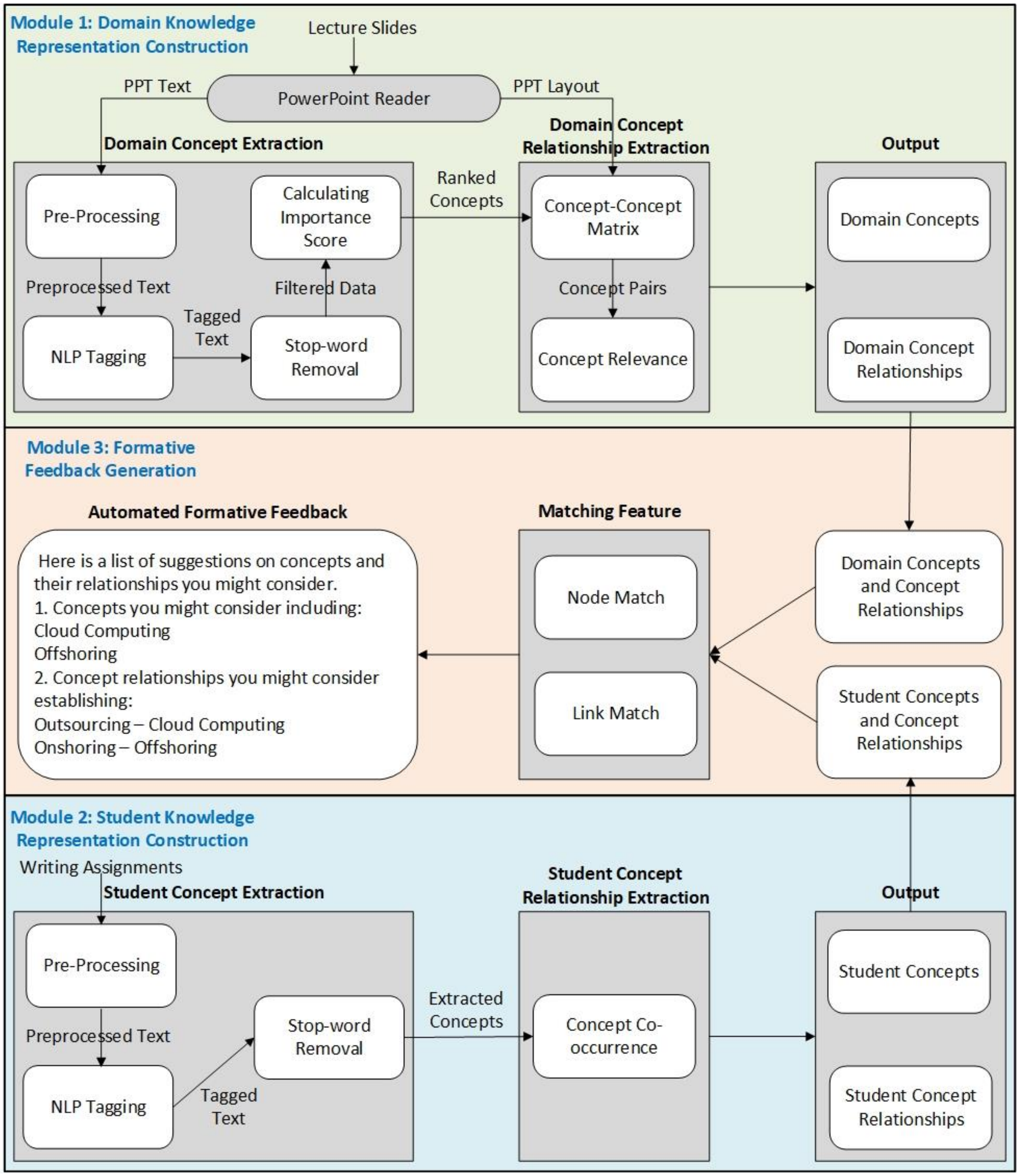

Figure 3.1 System architecture.

3.2 Module 1: Domain Knowledge Representation Construction

Formal lecture is a dominant teaching format within higher education, which is regularly supported with PowerPoint slides. In general, instructors (or domain experts) dedicate considerable efforts and expertise to produce semantically rich, semi-structured lecture slides based on extensive knowledge, experience, and relevant course materials (e.g., textbooks) (Atapattu et al., 2012). Previous studies (Atapattu et al., 2012, 2014a, 2014b) suggest that it is feasible to automatically construct domain knowledge representations from digital lecture slides for a specific course subject. In order to generate automated feedback in our study, domain concepts and the relationships among these domain concepts are identified from lecture slides, which are used as a reference to show the gap between actual and desired learning performance.

The module of domain knowledge representation construction is to extract domain concepts and domain concept relationships from lecture slides. For our research purposes, we consider the following as domain concepts: nouns, compound nouns, or noun phrases covered in instructors' lecture slides for a course subject. From the concept mapping point of view, a document can be formalized as a set of $D = \{C_d,\ R_d\}$, where $C_d = \{c_{1,}c_{2,\dots,}c_n\}$ is a set of all concepts and $R_d = \{r_{1,}r_{2,\dots,}r_n\}$ is a set of all concept relationships that can be extracted from the document (Zubrinic et al., 2012). Following prior study by Atapattu et al. (2012, 2014a, 2014b), syntactic parsing and Part-of-Speech (POS) tagging are implemented in our study to automatically extract the domain concepts and domain concept relationships from digital lecture slides for a specific course subject. The construction of domain knowledge representation in our research consists of two major steps: (1) domain concept extraction, and (2) domain concept relationship extraction.

3.2.1 Domain Concept Extraction

Domain concept extraction is the process of identifying the key concepts from the lecture slides, which consists of the following steps: (1) pre-processing; (2) NLP tagging; (3) stop-word removal; and (4) calculating importance score.

3.2.1.1 Preprocessing. To extract the domain concepts, the first step is to preprocess the text content of digital lecture slides of a given course as our knowledge source. In our approach, each bullet in the PowerPoint slide is equivalent to a sentence. If a bullet contains more than one sentence, it is considered as one sentence for analysis. The sentences with auxiliary information (bullet level and emphasized text) are taken as input. To process text content of Microsoft PowerPoint documents, the lecture slides are first converted into files in Rich Text Format (often abbreviated RTF), which can obtain the text content with rich text features such as title, bullet offset, and emphasized text (i.e., bold, italic, and underlined words). Then, all the RTF tags are removed, and the text content is extracted from the raw RTF files. All the text features are kept in order to help identify emphasized key concepts and concept relationships for our research purposes. To improve the concept extraction, the PowerPoint text content is pre-processed before preparing it for linguistic annotation. The pre-processing includes: (1) removing non-alphanumeric symbols; (2) processing punctuation marks; (3) correcting spelling errors; and (4) removing white spaces.

3.2.1.2 NLP Tagging. The Part-of-Speech (POS) tagger in the Stanford Core NLP project (Manning et al., 2014) is utilized to identify nouns and noun phrases in the text content. The POS tagging is the process of labeling each word in a sentence with its appropriate POS tag. The POS tagger can identify nouns (N), verbs (V), adjectives (ADJ), adverbs (ADV) and other POS definitions such as prepositions (P), conjunctions (CON), pronouns (PRO), and interjection (INT) in phrases or sentences. Compound nouns and noun phrases are then extracted as candidate concept terms.

3.2.1.3 Stop-Word/Phrase Removal. Some concept terms extracted in the previous step are not domain specific, such as "everything", "anyone", and "example". In the educational context, however, these terms are of no importance. In order to improve our results, a stop-word filter eliminates commonly used words that do not contain significance.

3.2.1.4 Scoring Concept Terms. This step is used to determine the importance of the extracted concept terms using a ranking model similar to Atapattu et al. (2012, 2014a, 2014b). In our modified ranking model, the following weighting factors are considered.

Concept Frequency

Concept frequency refers to the occurrence of each concept in the given lecture slides. As suggested in (Atapattu et al., 2012, 2014a, 2014b), log frequency weighting is assigned for each concept to normalize the occurrences within a controlled range. Thus, normalized concept frequency can still be an important factor in choosing important concepts while preventing a bias towards high frequency concepts in determining the threshold value for selecting important concepts.

$$\text{normalized concept frequency} = \text{Log2}(1 + \text{concept frequency}) \tag{3.1}$$

Concept Level

The PowerPoint layout (e.g., title, bullet point, and sub-bullet point) can help identify the level of each concept shown in the given lecture slide. In general, the concepts that occur in the title level are more important than the ones in the sub-bullet level. As a concept might appear more than once in different indentation levels, the average level of each concept is calculated. To do this, the level of each concept is identified and the summation of the level of each concept is calculated. For example, if there is a concept "outsourcing" that occurs once in different indentation levels, the average level of the concept "outsourcing" can be calculated by using the summation of the concept level divided by concept frequency. To normalize the concept level value and to ensure the higher the value the more important the term is, the maximum concept level for all concepts in the same PowerPoint slide set is used. In this study, based on the given datasets, the maximum concept level is 4. The average level of the concept is then calculated as below:

$$\begin{aligned}&\text{normalized average concept level}\\&\qquad = \text{maximum concept level} - \frac{\text{summation of concept level}}{\text{concept frequency}}\end{aligned} \tag{3.2}$$

Length of Concept

As suggested in related work (Ventura & Silva, 2012), concepts are units of knowledge made of words having some semantic meaning, and a compound concept that contains more than one word is more specific. In other words, the longer the concepts, the more specific they tend to be. For instance, the compound concept ("supply chain management") is more specific than the concept "management". Also, concepts in textual documents are usually described by noun phrases and these noun phrases are varied in length. In (Richardson et al., 2006), it is suggested that phrases with the lengths of 2, 3, 4, 5 words are common in concept maps. Thus, the length of concept plays a role in determining the importance of the extracted concepts and the number of words in each extracted concept is calculated in our approach.

Emphasized Concepts

Some concepts or terms are emphasized by instructors in the lecture slides to illustrate their importance in a course, frequently employing underlined, bold, or italic. As suggested in (Atapattu et al., 2012), these emphasized texts help choosing important concepts. The number of times each concept is emphasized is calculated in our approach.

Capitalized Concepts

Different from related work (Atapattu et al., 2012), captitalized concepts are also considered in our modified ranking model, because the initial letters in the concepts or terms are often capitalized for emphasis or when used as the headings of a lecture slide. In this approach, we consider whether there is a capitalized word in the concept and then calculate the number of capitalized words in the concept. For instance, in the phrase "Cloud Computing", there are two capitalized words in this concept.

Calculating Importance Score

Finally, the following weighting factors are included in our modified scoring model, including: (1) normalized concept frequency; (2) normalized average level of the concept; (3) the number of words in the concept; (4) number of times the concept is emphasized; and (5) number of capitalized words in the concept. The output of the scoring function will indicate how important a concept term is in the PowerPoint slides. As such, we call the output "Importance Score" for each concept and it is calculated using the formula below. By default, all these weights (w_n, where n = 1 to 5) are equal and they sum up to 1. In practice, the weights can be easily adjusted, and instructors can decide which factors should be assigned higher weights as needed.

$$\begin{gathered}\text{Importance Score (concept)} = \\ \text{w1} * [\text{normalized concept frequency}] + \text{w2} * [\text{normalizd average concept level}] \\ + \text{w3} * [\text{number of words}] + \text{w4} * [\text{number of times emphasized}] \\ + \text{w5} * [\text{number of capitalized words}]\end{gathered} \tag{3.3}$$

The top n ranked list of domain concepts $DC = \{dc_{1,}dc_{2,\dots,}dc_{\mathrm{n}}\}$ can be produced based on the Importance scores of all the extracted domain concepts, which indicate their importance in the given lecture slides. This is especially useful, when a student requests feedback before writing the assignment; in such a case, the top concepts can be provided as feedback.

3.2.2 Domain Concept Relationship Extraction

Domain concept relationship extraction is the process of establishing the relationships among domain concepts. To extract the domain concept relationships from lecture slides, PowerPoint layout as illustrated in Figure 3.2 is used as a key feature. Based on the information about different indentation levels of bullets (e.g., title, bullet, sub-bullet, and sub-sub-bullet) of the given lecture slides, the relationships among the extracted domain concepts can be identified. As part of the output of the domain knowledge representation construction, each concept is given an ID and the ID is recorded in the concept hash table. The concept-to-concept relationship is recorded in an adjacency matrix, where each cell C_map[*i*][*j*] = 1 means that concept *j* and concept *i* have direct hieratical relationship. For example, the concept C1 is extracted from the first level. The concepts C2 and C3 are extracted from the second level. There are direct hieratical relationships among these concepts: C1 and C2, C1 and C3.

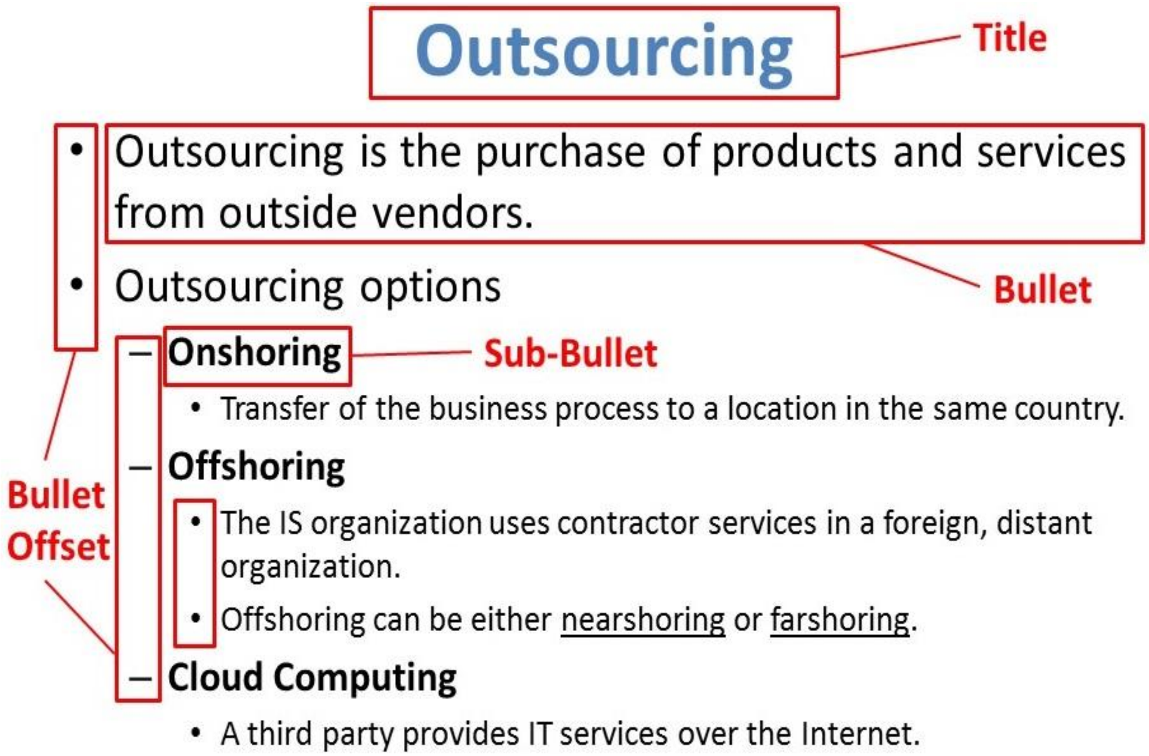

Figure 3.2 Example of PowerPoint layout.

3.2.3 Concept Distance Calculation

Subsequently, by using Floyd's algorithm (1962), the shortest distance between two concepts is computed and recorded in the concept-distance matrix as illustrated in Figure 3.3 as well. The pre-calculated shortest distance between each term pair in the same PowerPoint slide set will improve the efficiency during real time generation of concept terms as feedback in Module 3.

	C1	C2	C3	C4	C5
C1	-				
C2	(1, 1)	-			
C3	(1, 1)	(0, 2)	-		
C4	(1, 1)	(0, 2)	(0, 2)	-	
C5	(0, 2)	(0, 3)	(0, 3)	(1, 1)	-

Figure 3.3 Example of concept-distance matrix.

3.3 Module 2: Student Knowledge Representation Construction

The module of student knowledge representation construction is to extract student concepts and student concept relationships from students' writing assignments. Different WTL assignments, such as summaries, reflective essays, online discussions, etc., have been widely used by instructors to help students understand and apply academic content or conceptual knowledge presented in a course. In our system, these textual assignments are used as the data source to represent an individual student's current state of knowledge and construct student knowledge representation. With the writing assignments of individual students as input, the module can automatically extract the concepts and their relationships from them, and then produce the lists of student concepts and student concept relationships.

3.3.1 Student Concept Extraction

Following most of the same steps for extracting domain concepts from lecture slides, to extract the concepts from students' written text, the student concept extraction includes one extra step -- splitting statements at the occurrences of periods. This new step is needed in this module to preprocess concepts from a student assignment, because unlike PowerPoint slides, a typical assignment is composed of complete sentences instead of bullets. Then, nouns and noun phrases in students' written text are extracted using the same approach for identifying domain concepts. Also, the stop words without information values are removed. By doing so, the set of student concepts $SC = \{sc_{1,}sc_{2,\ldots,}sc_n\}$ can be extracted from individual students' written text.

3.3.2 Student Concept Relationship Extraction

To identify the relationships among student concepts, a student assignment is separated into paragraphs in our system. For student concepts extracted in the same paragraph, they are treated as associated concepts for our research purposes. In other words, if the noun or noun phrases as concepts are written in the same paragraph, they are considered as associated with each other. By doing so, the student concept relationships $SR = \{sr_{1,}sr_{2,\ldots,}sr_{\mathrm{n}}\}$ can be produced. The concept-to-concept relationship is recorded in an adjacency matrix, where each cell C_map[i][*j*] = 1 means that concept *i* and concept *j* occur in the same paragraph.

3.4 Module 3: Formative Feedback Generation

3.4.1 Design Considerations

Meaningful learning takes place by assimilating new concepts and propositions into existing cognitive structures held by learners (Novak & Canas, 2007). As the formative feedback provided in the system aims to support meaningful learning of conceptual knowledge based on what a student has or has not already written, there is a need to compare the differences between domain knowledge representation and student knowledge representation. The module of formative feedback generation is the core of the system, which compares the two lists of concepts and concept relationships extracted from the given lecture slides and a writing assignment, and produces a list of suggestions as automated formative feedback that shows the missing concepts or unestablished relationships among concepts that students might consider for revisions.

In the proposed system, formative feedback is provided in the form of facilitative feedback, which consists of two major components: (1) missing concepts: a list of concepts that are not included in an individual student's writing assignment; and (2) unestablished concept relationships: a list of missing concepts that are associated with the concepts already written in an individual student's writing assignment. For instance, if any concept or concept relationship is missing, formative feedback is generated as follows: "*Suggested concepts you might consider including: Cloud computing; Suggested concept relationships you might consider establishing: Outsourcing – Cloud computing...*"as shown in Figure 3.4. As the automated feedback is tailored for a specific assignment draft, students can revise their assignments by assimilating the most important and relevant domain concepts into their existing knowledge structure to construct a more integrated knowledge structure. By doing so, students are expected to learn the key course concepts more meaningfully during a WTL activity.

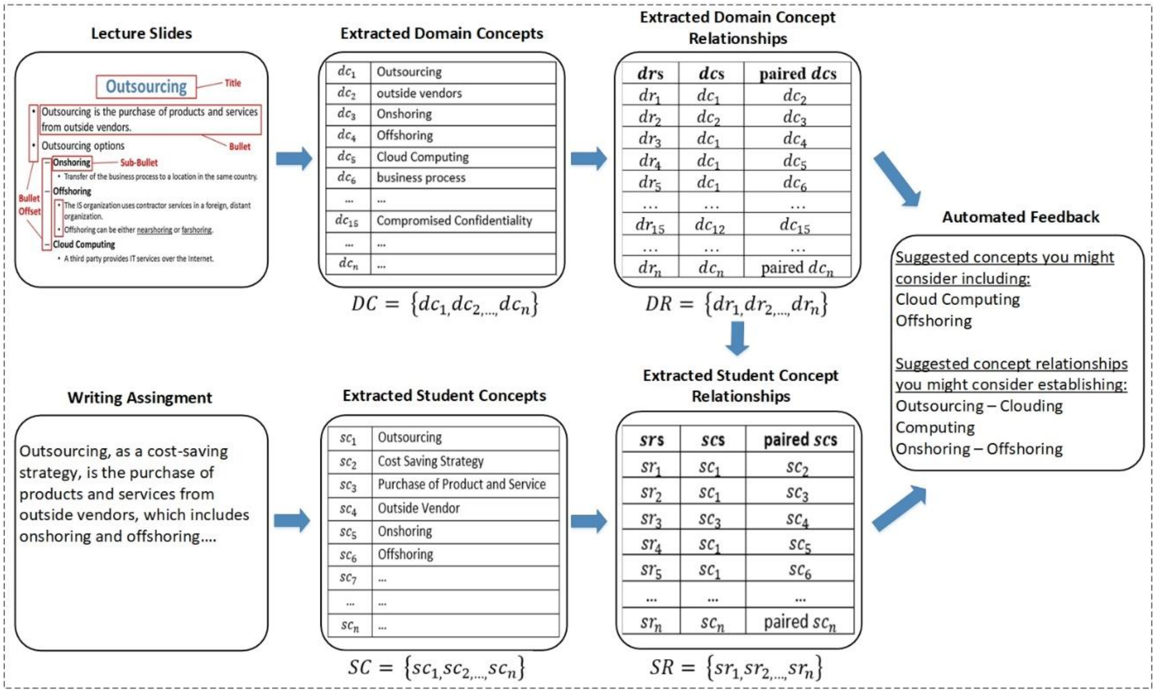

Figure 3.4 Formative feedback generation.

3.4.2 Generating Recommendations

Once the domain knowledge and student knowledge are represented, the first task is to map a student's knowledge representation to the domain knowledge representation which can be implemented as a task comparing the two knowledge representations: (1) identifying matching concepts, i.e., identifying concepts appearing in both the student concept set {SC} ***and*** also the domain concept set {DC}; and (2) identifying matching concept relationships, i.e., identifying those concept relationships appearing in both the student concept relationship set {SR} ***and*** also in the domain concept relationship set {DR} with identical connecting concepts.

Every time a student requests feedback after he or she writes at least a sentence, based on whether there is any matching concept, there are two scenarios, and each will have a corresponding method for feedback generation.

3.4.2.1 Scenario 1: No Matching Concept. If there is no matching concept, a ranked list of domain concepts, based on their importance scores (calculated using Formula 3.3 above), are provided as feedback for the student to consider for inclusion. At this stage, the most important goal is to offer the student key concepts from the lecture as starting points. No concept relationship is recommended in this scenario.

3.4.2.2 Scenario 2: At Least One Matching Concept.

3.4.2.2.1 Step 1: Recommending Relevant Important Concepts for Inclusion.

In such a case, it is desirable for the suggested concepts for inclusion to be as closely related to ***all*** the matching concepts in the student's assignment as possible, instead of just an individual matching concept term. For example, if a student writes "information systems" and "information technology" in the assignment and the two terms both appear in the domain concept set, they are considered matching concepts and the suggested concepts generated by the system must be as closely related to both of these terms as possible.

To achieve this goal, the assignment is represented as a collection of matching concepts, which can range from one to many matching concepts. Then the ***average*** distance between any domain concept and the matching concepts extracted from the student assignment can be quickly determined by using the pre-calculated domain concept distance described in Section 3.2.3 to improve the efficiency of the system. The average distance yields outputs where the farther the distance between two concepts, the more distant the relationship between them. Therefore, in order for the larger outputs to indicate a higher relevance, a conversion function $f(x)$ was used. The complete Relevance formula is listed below. $\boldsymbol{\lambda} = 0.5$ by default.

$$\text{Relevance (concept, assignment)} = f\text{ (Average (Distance (concept, }i\text{th matching concept in the assignment)))} \quad (3.4)$$

$$\text{where } f\text{ (x)} = \lambda * \exp(-\lambda * \text{x})$$

As described previously, the recommended concepts for inclusion should take into account both the *importance* (how important a concept term is in the PowerPoint slides, generated using Importance score Formula 3.3) and *relevance* (how relevant a domain concept term is to all the student concept terms in the assignment, generated using Relevance score Formula 3.4). Formula 3.5 is designed to reflect such a consideration, where the Relative score of a concept term regarding an assignment draft is the product of Importance score and Relevance score.

$$\text{Relative Score (concept, assignmnet)} = \text{Importance Score (concept)} \times \text{Relevance (concept, assingment)} \quad (3.5)$$

The top domain concepts with the highest Relative scores are recommended for inclusion.

3.4.2.2.2 Step Two: Recommending Concept Relationships for Inclusion.

All the matching concepts' corresponding concept relationships are retrieved from {DR} and compared with those in {SR}; those retrieved but not yet in {SR} will become candidate concept relationships for recommendation. The sequence of recommendation will be based on the *relative scores* of the ***associated concepts*** calculated based on Formula 3.5. In other words, similar to how concepts are recommended for inclusion, concept relationships are also recommended based on what matching concepts and matching concept relationships have already been written in the assignment. As a result, closely related concept relationships will be recommended first.

For example, two matching concepts “Information Systems” and “Information Technology” and their domain concept relationships: {“Information Systems”- “Systems Analysis and Design”}, {“Information Systems”-“Technology Acceptance Model”}, {“Information Technology”-“Information Systems”}, and {“Information Technology”-“Network Technology”}. Because “Information Systems” and “Information Technology” are already identified as matching concepts in the student’s writing assignment, the system will not recommend these two concepts for inclusion. If these two concepts occur in the same paragraph of the assignment, the system will not recommend the student to establish the relationship between these two concepts. The other three associated terms will be recommended to form concept relationships with the existing matching concepts in the assignment, based on their average distance to “Information Systems” and “Information Technology”. Suppose “Systems Analysis and Design” has the highest relative score, the system will recommend the student to establish a relationship between “Information Systems” and “Systems Analysis and Design” first.

3.5 System Implementation

This study implemented the automated feedback generation approach in a prototype system that is applicable as a real-time operation. The proposed system is a writing and learning support system that provides automated, actionable, and individualized formative feedback through the comparison between the knowledge representation of lecture slides and that of a writing assignment to improve student learning in WTL activities. The system automatically performs concept extraction from lecture slides, analyzes the importance of course concepts, and identifies the associated relationships among these concepts.

Our proposed system focuses on the support and promotion of meaningful learning through formative feedback in WTL activities, which refer to writing assignments or tasks that help students think through core concepts or ideas presented in a course (Forsman, 1985). Using the Axure RP platform, a prototype of the user interface was first developed following an iterative development cycle. Afterwards, the prototype system was implemented with Django (i.e., a Python web server framework). Figure 3.5 shows the first screen of the interface of the proposed system. This section describes the system user interface and illustrates its key functions. Protocol analysis was conducted and will be discussed in Section 4.2 in the next chapter.

Figure 3.5 System interface: log in.

After logging into the system, students can select the assignment to write as illustrated in Figure 3.6.

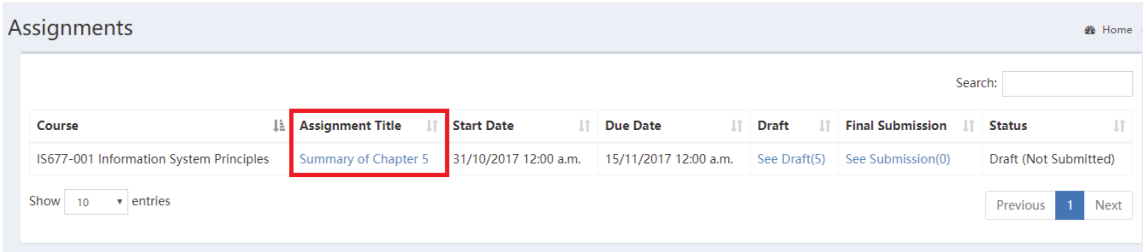

Figure 3.6 System interface: select writing assignment.

After reading the assignment instructions as illustrated in Figure 3.7, students can click the “Write Assignment” button and start to write the assignment.

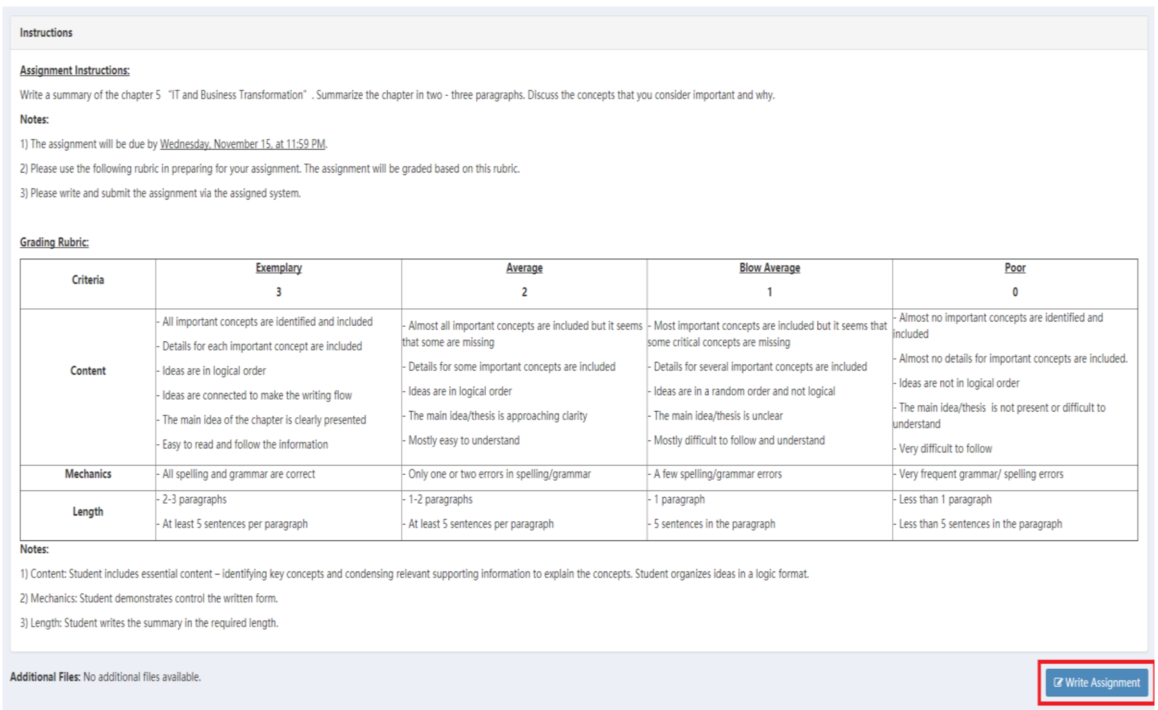

Figure 3.7 System interface: read assignment instructions.

As shown in Figure 3.8, students can work on the assignments from any computer with an Internet connection and a compatible browser.

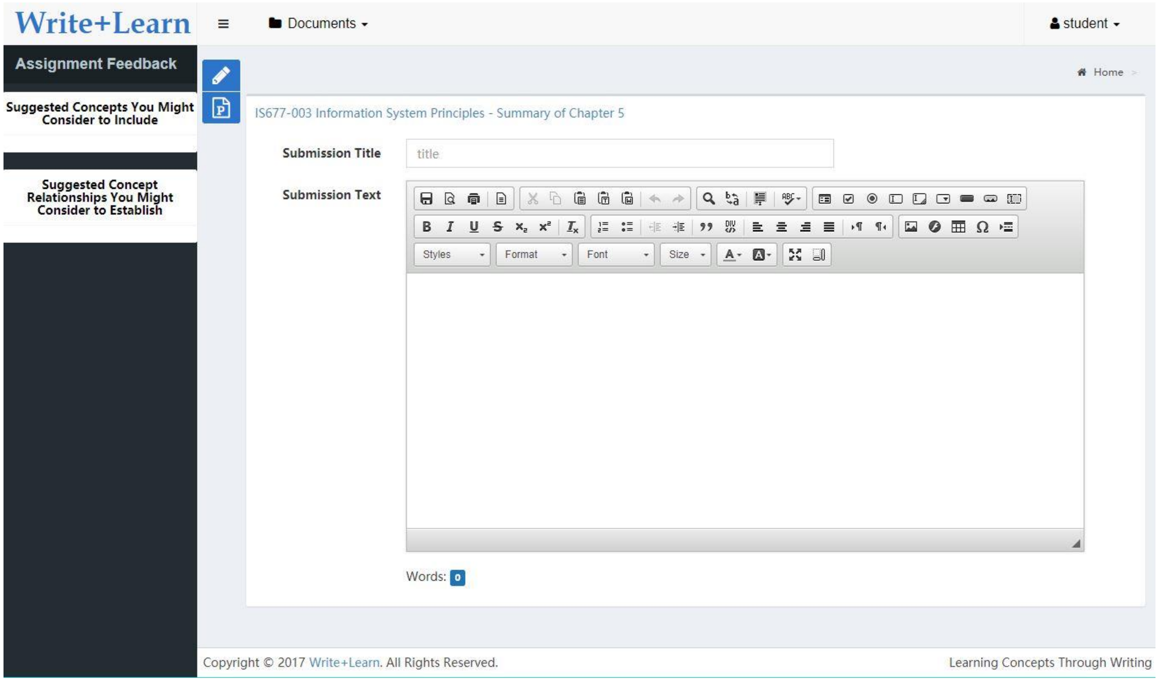

Figure 3.8 System interface: write assignment.

To support student learning and writing, the system allows for multiple revisions and editing. As illustrated in Figure 3.9, in interactive mode, students can ask the system to generate formative feedback by clicking the "Request Feedback" button at any time, either in the middle of or after having completed an assignment draft. This enables feedback to be immediately actionable and motivates students to make further revisions of their assignment drafts based on the suggestions.

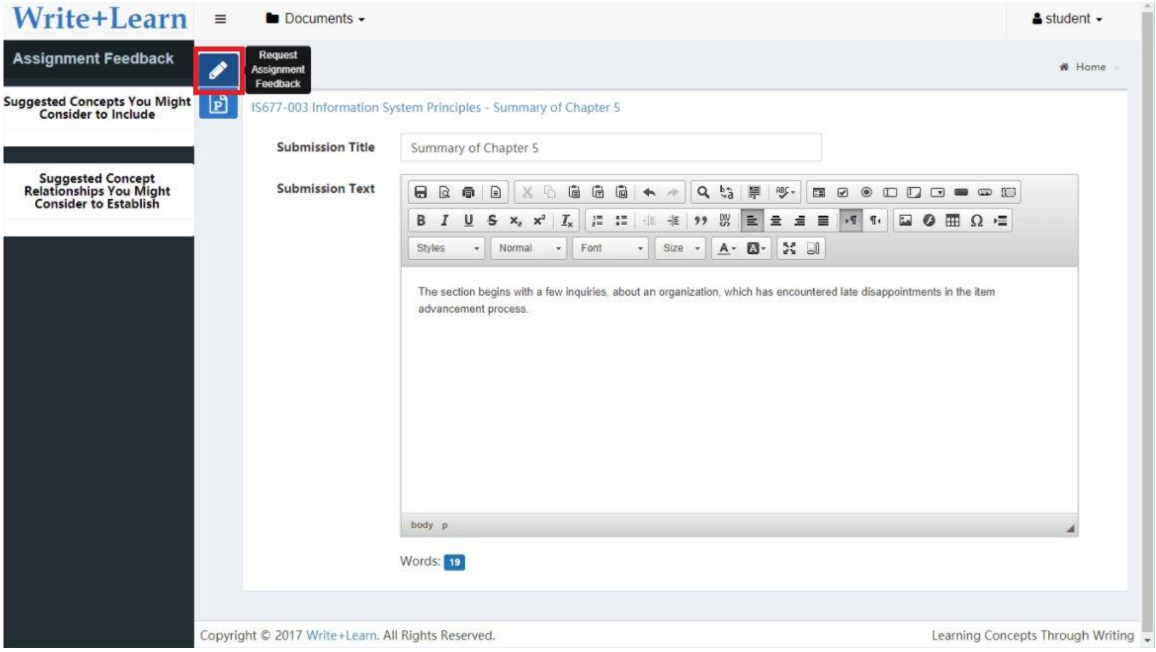

Figure 3.9 System interface: request formative feedback.

As illustrated in Figure 3.10, the sidebar shows a list of suggestions that include the missing concepts or unestablished concept relationships based on what is included in the input assignment draft. As suggested in (Kintsch et al., 2000), it is easy to overwhelm users and confuse them with the rich feedback the system can provide. In the system, only the top 10 suggestions, including missing concepts and unestablished concept relationships, are displayed in the first page. If the students would like to view more feedback, they can go to the next pages for more suggestions. In our studies, the system could display up to the top 50 suggestions including missing concepts and unestablished concept relationships respectively in a default setting. In practice, instructors can have the option to select the number of suggestions to be presented in the system. For the scope and purpose of this research, we only consider the existence of a relationship between concepts. Following the suggestion in related work (Leake, 2006), we do not provide information like labels about the relationship between two concepts. In WTL activities, students are expected to apply the key domain concepts in a writing assignment. It is more meaningful for them to associate the recommended concepts and concept relationships with what is already written. There is no need for students to consider establishing the relationship labels like concept mapping tasks during the process of writing an assignment.

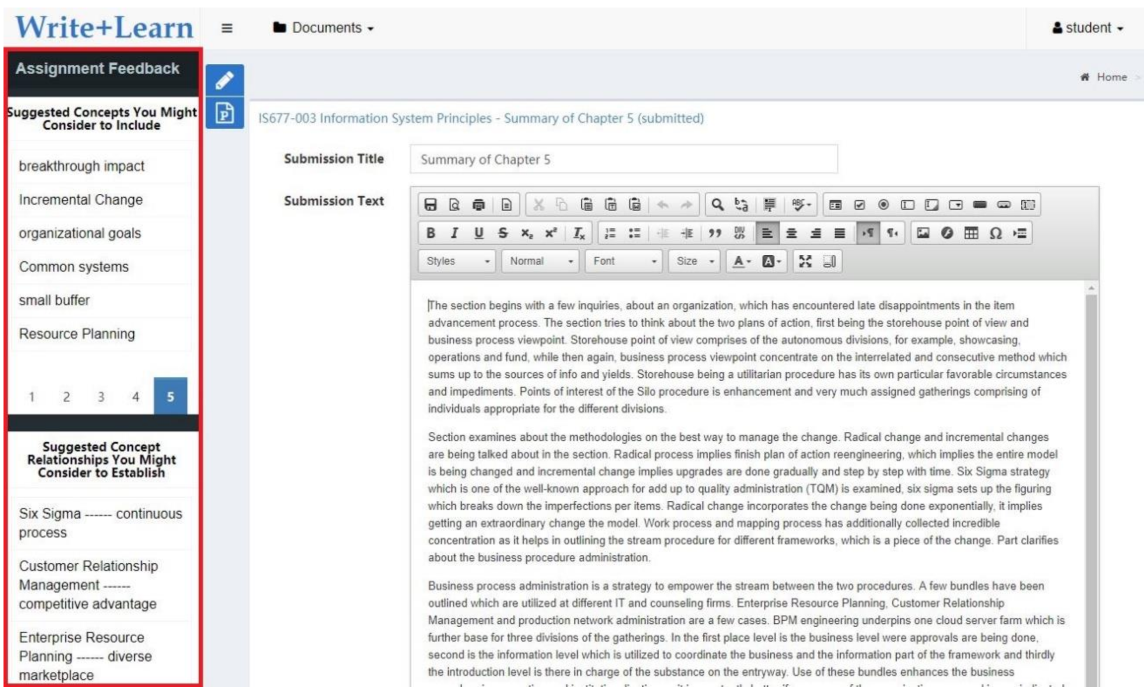

Figure 3.10 System interface: review formative feedback.

The system allows students to make revisions based on the feedback given to the assignment draft by clicking the "Edit Submission" button. Finally, students can submit their work by clicking the "Submit Assignment" button as illustrated in Figure 3.11.

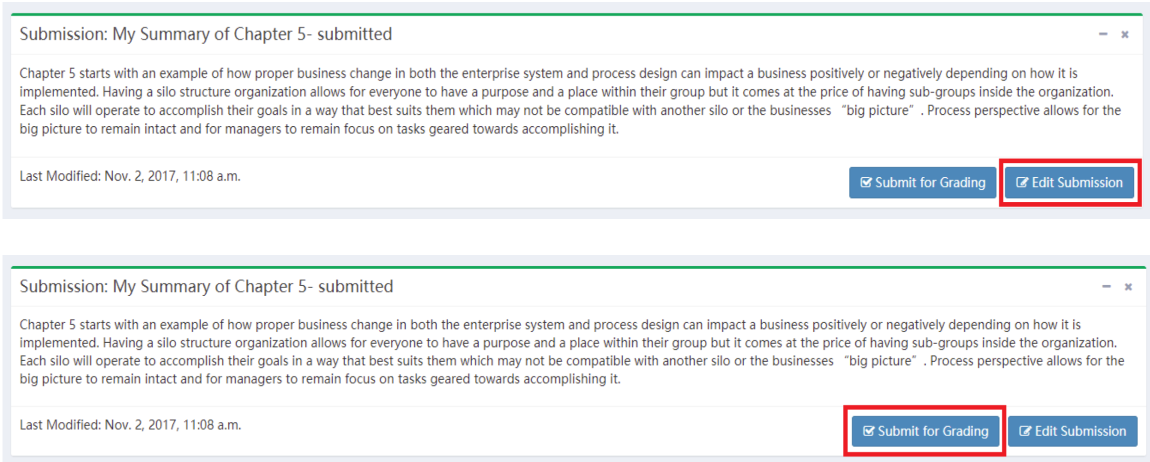

Figure 3.11 System interface: revise or submit assignment.

Students can get access to previous revisions of assignment drafts to view their progress or refer to the corresponding lecture slides to reflect on their learning. In the proposed knowledge-centered and learner-centered environment, these scaffolds can improve their conceptual development, reflect on their knowledge gap, and demonstrate their learning progress via different assignment drafts in WTL activities.

3.6 Summary

This chapter discusses the architecture and implementation of the proposed system, as well as illustrates how the automated feedback can be generated and how students request and use the feedback in a WTL activity. The proposed system consists of three major modules: (1) domain knowledge representation construction; (2) student knowledge representation construction; and (3) formative feedback generation. First, all the concepts can be extracted from the lecture slides and writing assignments with NLP techniques. To refine the results, a modified scoring model was employed to ensure that all the extracted concepts can be ranked based on their importance. Second, as for concept relationship extraction, PowerPoint layout such as different indentation levels of the given lecture slides can be used as a key feature to define the relationships among concepts. By using Floyd's algorithm (1962), the most relevant concepts of any given concept can be identified. As the output of domain concept representation, the list of domain concepts and concept relationships can be identified based on both importance and relevance. Finally, based on the comparison of two knowledge representations, actionable and individualized formative feedback can be automatically generated.

CHAPTER 4

PRELIMINARY STUDIES

In this research, the automatic approach to generate formative feedback based on knowledge representations constructed from instructors' lecture slides and students' writing assignments is proposed. The purpose of developing the proposed prototype system is to support student learning of conceptual knowledge during the processing of writing an assignment without increasing the instructors' workload. To achieve our research goals and refine the design of the main study, the following preliminary studies were conducted: (1) a study to evaluate the quality of extracted domain concepts; (2) a think-aloud protocol analysis to explore students' perception of the prototype system; and (3) a pilot study to examine the system performance and test study instruments to be used in the larger-scale studies. In this chapter, the results of these preliminary studies are discussed.

4.1 Quality of Extracted Domain Concepts

4.1.1 Dataset

To evaluate the quality of extracted domain concepts, we selected an introductory graduate level Information Systems course at New Jersey Institute of Technology (NJIT) to conduct this experiment. Because all the courses at NJIT are coordinated, instructors of the different sections of the same course use the same course materials including the lecture slides. After working with the course coordinator, we selected two required writing assignments to be used for this study. The lecture slides of two chapters from the course textbook associated with the writing assignments used in this study were utilized by the system to extract key concepts.

4.1.2 Method

Two independent domain experts were invited to participate in the evaluation. One of the evaluators was the instructor who has been teaching this course for many years and the other was an information systems doctoral candidate who was a teaching assistant for this course. Both of them were experts in the course content with extensive experience in tutoring and assignment grading. The tasks of each evaluator included: (1) marking the key concepts extracted by the proposed system as 1 (a key concept) or 0 (not a key concept); and (2) identifying the key concepts from two chapters of lecture slides respectively. The evaluation can reflect upon the quality of extracted concepts based on their knowledge and perception. The evaluators received the information about the purpose of the study and the use of extracted concepts in the proposed system before the study.

The system's ability to extract key concepts from instructors' lecture slides and students' writing assignments is of critical importance before any feedback can be generated, and it is evaluated by Precision and Recall measures, which are two standard measures from information retrieval. Precision is defined here as the proportion of system extracted concepts that are considered important by the domain experts. Recall is defined here as the proportion of key concepts identified by the domain experts from lecture slides that are also extracted by the system.

$$\text{Precision} = \frac{|\{\text{Human Identified Key Concepts}\} \cap \{\text{Machine Extracted Key Concepts}\}|}{|\{\text{Machine Extracted Key Concepts}\}|} \quad (4.1)$$

$$\text{Recall} = \frac{|\{\text{Human Identified Key Concepts}\} \cap \{\text{Machine Extracted Key Concepts}\}|}{|\{\text{Human Identified Key Concepts}\}|} \quad (4.2)$$

4.1.3 Results

Table 4.1 summarizes the evaluation results of the quality of extracted domain concepts in terms of Precision. Regarding Precision, a few noun or noun phrases automatically extracted from the lecture slides are too generic, such as “key aspects” and “web browser”. Although the stop-word removal was implemented in concept extraction, the system still generated a few generic terms that carried little domain-specific information and accordingly were not considered as key concepts by the evaluators. Thus, there is a need to extend our stop word/phrase removal to exclude these generic terms.

Table 4.1 Quality of Extracted Domain Concepts Evaluation Results - Precision

	Evaluator 1		Evaluator 2		Evaluator 1 ∩ Evaluator 2		Match		Percentage of Agreement (%)	
Dataset	One	Two	One	Two	One	Two	One	Two	One	Two
Number of Machine Extracted Key Concepts	50	70	50	70						
Number of Machine Extracted Key Concepts Also Identified by Human Evaluators	39	49	40	50	35	43	41	53	82	76
Precision (%)	78	70	80	71						

Regarding Recall, other than the separate results based on different experts' inputs, we also provide here the Recall of key concepts that were chosen by both evaluators that were also extracted by the system (93% and 91%) as listed in Table 4.2. In other words, we believe concepts chosen by both experts were truly the most important ones, so the system's ability in identifying them is even more important.

Table 4.2 Quality of Extracted Domain Concepts Evaluation Results - Recall

	Evaluator 1		Evaluator 2		Evaluator 1 ∩ Evaluator 2	
Dataset	One	Two	One	Two	One	Two
Number of Human Identified Key Concepts	18	14	16	15	14	11
Number of Human Identified Key Concepts Also Extracted by Machine	15	11	14	12	13	10
Recall (%)	83	79	88	80	93	91

Overall, our results are comparable to those in related studies (Aguiar et al., 2016), where the reported precision and recall on the analysis of the automatically identified concepts from texts are only 47% and 67%, respectively. Choosing key concepts is subjective; some concepts that are considered as important by one evaluator might be considered not as important by another evaluator. For instance, "Six Sigma" is assessed by the second evaluator as a key concept; however, the first evaluator did not identify it as a key concept from the lecture slides. As summarized in Table 4.1, the percentages of agreement between two evaluators are 82% for the first dataset and 76% for the second dataset, which indicate an acceptable agreement in concept extraction. As a rule of thumb suggested in previous studies (Hartmann, 1977; Stemler, 2004), when using percentage of absolute agreement, values from 75% to 90% demonstrate an acceptable level of agreement.

We did not invite the experts to evaluate the suggested feedback as those were generated specifically for a student assignment draft and there were too many for experts to review. Thus, we relied on the survey of the participants to inform us the quality of the generated feedback and results are discussed in the next chapter. As stated in (Atapattu et al., 2014a), it is practically challenging for machines to outperform humans in a corpus like lecture slides since there is no well-defined structure for writing course materials. Our experimetal results are comparable with those reported in similar studies. As the next step of this research, further user studies were conducted to test the system performance and experimental design for the main study.

4.2 Protocol Analysis

To evaluate the usability of the proposed system, think-aloud protocol (Boren & Ramey, 2000) was used to understand students' perceptions of the prototype system used in our study. The think-aloud protocol analysis was conducted to observe and record students performing a mock assignment using the system in a computer laboratory environment.

4.2.1 Method

For this study, a group of five university students (N=5; 3 female and 2 male) were recruited via word of mouth and in-class announcements during the fall semester of 2018 at NJIT. All were graduate students between 21 and 24 years old. Four of them were studying in the field of Information Systems and one student was majoring in Engineering Management. Each usability testing session, approximately 35-45 minutes, was audio taped with the consent of the participant. Usability testing sessions were scheduled at a time that was convenient for each individual participant. Each student was offered $10 for participating in the session. None of the subjects had any former experience with think-aloud studies and all of them had substantial internet experience. Each student was informed of his or her rights as a research study participant and was then asked to sign a consent form before starting the usability testing session.

Upon completion of the consent form, the participants were given basic instructions describing the research task, the Write-and-Learn prototype system, and the format of the think-aloud protocol. After that, the participants were given a computer equipped with Camtasia Replay to record the participant's on-screen activities and utterances. The participants were given about 20 minutes to complete the mock assignment. The mock assignment required the participants to respond to the following question: "*What is IT business transformation? Please complete the question in two or three paragraphs.*" The question was provided on the printed instructions and was also made available in the system. The participants were asked to work on the writing task as they normally would, except that they must think out loud as they completed the writing task. During the task, we limited the interaction with the participants, only intervening when necessary such as reminding the participants to verbalize their thoughts while performing the writing task.

Data collected through the think-aloud protocol analysis study included: (1) the spoken thoughts and utterances of the participants which were transcribed for analysis; and (2) the system input from each participant. Think-aloud protocols captured a detailed record of what was going on in the subject's mind during the process of using the system to complete the mock assignment. Upon completion of the think-aloud protocol, a brief post-activity interview was conducted to describe retrospectively how they used the system to complete the writing task. This interview prompted participants for immediate reactions to their assigned system, the writing task and instructions, and the perceived value of the provided automated feedback.

4.2.2 Results

The think-aloud study allowed a better understanding of how the prototype system functioned and helped examine the processes that student subjects used as they performed a mock assignment. Overall, the participants provided positive feedback. The verbal protocol data revealed that the participants had no difficulty in finding the writing assignment and requesting feedback. All participants commented that the main navigational links were obvious to the user and no changes needed. The system was functioning as intended and the participants found nothing confusing or disliked. All of them liked the design of the system and thought it was easy to understand and user-friendly. For example, it was stated that "*The UI is good. It is very easy to understand. Everything is very easy.*"

When asked about what they liked most about the system, participants commented that "*the system gave you suggestions, which was good.*" "*As compared to Moodle, it gave you a lot more. I think this is better as it gives you suggestions. They can make your grade better*." Other students said, "*Sometimes we perceive something else is needed in the assignment, but we do not know what the professor and the grading person are expecting. You know their expectation and what keywords we should be using. Sometimes we just use any words…If you can give me feedback, I feel it is very positive for someone who really want their scores well.*" One comment was related to a participant's feeling about receiving feedback from their instructor in general: "*The professor or TAs only gave you feedback after the assignment was done or the grading was done. There is no way for us to know what we perceived was right or wrong. There is no direction. We do not know what they are expecting…I just need to know what topics need to be studied, what are the things should be talked about and included in my assignment. In that matter, that would be very useful for students*."

The participants were asked to request the feedback and revise the assignment further after reviewing the feedback. In terms of the perceived value of feedback, participants provided the following comments: "*That is very useful. Many times, when we write assignments, we do not know what keywords we are missing. So, the keywords can help us get better grades. I think it is really good.*" "*They (the feedback) were accurate, because they are automatically taken from the slides. They are not random words, which means these words have some value. They are not randomly put. They were relevant.*" However, only one student mentioned that there is no need to get so much feedback, as "*I do not want to go through the long list. Now I can see five pages, but two pages would be enough.*" As one of the default settings, the system displays top 50 suggestions in five pages (10 suggestions per page), including missing concepts and unestablished concept relationships, respectively. In practice, instructors can easily adjust the number of suggestions to be presented in the system.

In short, student participants evaluated the system's clarity, simplicity, navigation and features as positive. From their points of view, the system was intuitive, easy to use, easy to learn and user friendly; they mostly liked the idea of receiving the feedback during the process of writing an assignment. Their responses confirmed that the automated feedback system can provide students with novel forms of assistance in their reflection and learning of course concepts in a WTL activity. Throughout the studies, the system worked well, and no severe technical problems were found.

4.3 Pilot Study

A pilot study was conducted to preliminarily examine the study instruments used in this research prior to beginning the main study. To adequately test the experimental design in a classroom setting, two sections of students in an Information Systems graduate course (i.e., IS 677 Information Systems Principles) were invited to participate in the pilot study during the fall semester of the 2017 academic year at NJIT.

4.3.1 Theoretical Framework

Given the widespread use of technology in education, it is of significant concern how and why users accept and use the technology (Jiang & North-Samardzic, 2015). It is suggested that any new technologies and systems could fail because the end users do not accept using them (Al-Assaf et al., 2015). There are many theories of technology acceptance used to gauge the perceptions of end users, such as the Technology Acceptance Model (TAM) (Davis, 1989). As one of the most robust and influential models on the adoption behavior of information technology, TAM focuses on predicting information system acceptance and diagnosing design problems for a new system (Szajna, 1996).There are two major factors that TAM depends on to predict user acceptance of any technology: perceived usefulness and perceived ease of use.

Based on TAM and other related models of user acceptance, the Unified Theory of Acceptance and Use of Technology (UTAUT) model was developed by Venkatesh et al.

(2003) to comprehensively predict individual acceptance of information technology, i.e., intention and behavior. Prior studies (Attuquayefio & Addo, 2014; Debuse et al., 2008; Jiang & North-Samardzic, 2015; Liao et al., 2004; Marchewka & Kostiwa, 2007; Mathur, 2011) indicate that the UTAUT model can be used to evaluate the users' acceptance of web-based learning environments. UTAUT identifies four constructs as direct determinants of user acceptance and usage behavior: (1) performance expectancy: the degree to which an individual believes that the use of the system can help achieve gains in job performance; (2) effort expectancy: the degree of ease associated with using the system; (3) social influence: the degree to which an individual perceives how important others believe it is to use the system; and (4) facilitating conditions: the factors that influence an individual's belief to perform a procedure. Additionally, four moderators (i.e., age, gender, experience, and voluntariness of use) are found to improve the predictive ability of the UTAUT model in previous studies (Venkatesh et al., 2016).

Based on a review of previous research, a variation of the UTAUT model is used in this research as illustrated in Figure 4.1, which includes the following constructs: self-efficacy, effort expectancy, facilitating conditions, perceived feedback effectiveness, perceived enjoyment, perceived ease of use, perceived usefulness, perceived learning and recommendation for use. Validated constructs adopted from Ventkatesh et al. (2003) are operationalized such that the wording is changed to account for the context of our research study. To adapt the original UTAUT model to the educational context, several variables have been added to ensure a better understanding of technology acceptance by students, such as self-efficacy, perceived enjoyment, and perceived learning suggested by related work (Marchewka & Kostiwa, 2007; Oye et al., 2012; Terzis & Economides, 2011). Social influence is not incorporated in our model, because this is a first use of a brand-new system for the subjects. Also, the moderators (e.g., age, gender, experience, and voluntariness of use) are not explicitly included from the original UTAUT model, because most of the factors do not vary sufficiently in this research. Also, the variable “voluntariness of use” becomes an issue as the system must be used in the study.

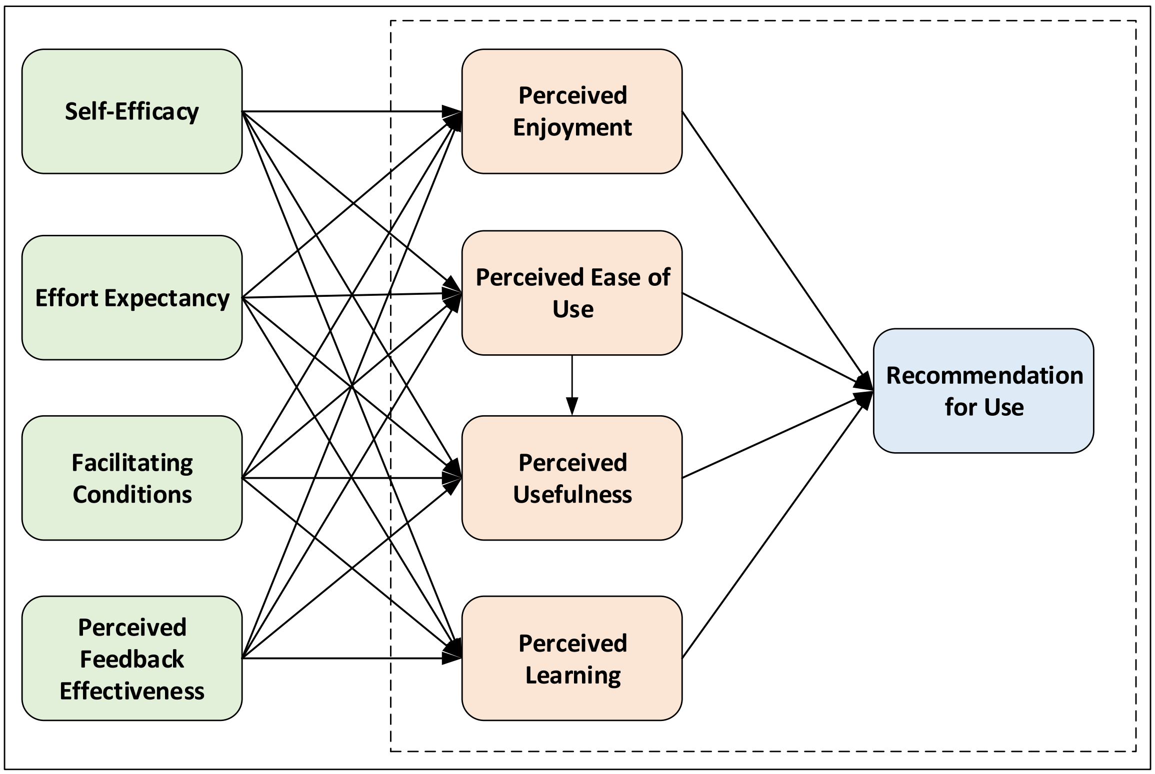

Figure 4.1 Variation of the UTAUT model.

The survey instrument was designed based on constructs validated in prior research (Davis, 1989; Venkatesh et al., 2003) and adapted to the context of this research. This study incorporated a total of forty measurement items on the questionnaires to measure nine variables, which included six items for self-efficacy, two items for effort expectancy, five items for facilitating conditions, four items for perceived feedback effectiveness, four items for perceived enjoyment, five items for perceived ease of use, four items for perceived usefulness, nine items for perceived learning, and one item for recommendation for use. All questionnaire items were measured using a five-point Likert scale. Appendix D lists the definition and the measurement items for each variable. Complete copies of all surveys for this study can be found in the Appendix C. The reliability analysis was conducted to measure the internal validity and consistency of questions used for each construct by calculating Cronbach's alpha coefficient. It is suggested that a Cronbach's alpha of 0.6 and above was considered reliable when judging a scale. In this study, the reliability of the Likert-type items in the questionnaire was measured to be 0.761 implying that the instrument was reliable.

In this study, the independent variable "Self-Efficacy (SE)" refers to the belief that one has the capability to perform a particular behavior, including system-specific self-efficacy and writing self-efficacy, which refers to the individual's perceptions of his or her capacity to use the system to write the assignment. The independent variable, "Effort Expectancy (EE)" is defined as a person's perceived course difficulty and assignment complexity in our study context. The independent variable, "Facilitating Conditions (FC)" refers to the perceived enablers or barriers in the environment that influence a person's perception of ease or difficulty of performing a task, including the clarity, perceived quality, and fairness of the assignment, grading rubrics and procedures. The independent variable "Perceived Feedback Effectiveness (PFE)" relates to student perception of the quality and effectiveness of the provided feedback.

The intervening variable “Perceived Enjoyment (PE)” refers to the degree to which a person believes that using a system is enjoyable, which reflects the concept of intrinsic motivation. As described in related work (Wu & Hiltz, 2004), perceived enjoyment is part of perceived motivation that has a significant impact on student learning. The intervening variable “Perceived Ease of Use (PEU)” refers to the degree to which a person believes that using the system is free from effort. The intervening variable “Perceived Usefulness (PU)” refers to the degree to which a person believes that using the system enhances his or her academic performance. “Perceived Learning (PL)” is defined as the degree to which a person believes that using a particular system improves his or her learning, which was operationalized through questions in the post-assignment surveys to capture student perceptions of their learning. Finally, the dependent variable “Recommendation for Use (RU)” is substituted for “Intention to Use”, which refers to the degree to which a person would recommend the system to be used in the future.

4.3.2 Method

In the pilot study, two sections of students (with a total of 32 students) were randomly assigned into one of the two groups: one experimental group (N=16) doing a writing assignment with automated feedback and one control group (N=16) doing the same writing assignment without automated feedback. Participants in the experimental and control groups were taught by the same instructor. With the instructor’s approval, the research activities lasted four weeks, including an extra credit writing assignment and two online survey questionnaires. Student participants were informed that their participation in the study was completely voluntary but would result in extra credit. Students who did not wish to participate in the research were offered an equivalent alternative assignment to obtain the same amount of extra credit. Upon completion of the consent forms, the participants were given basic instructions describing the research tasks, including pre-assignment survey, writing assignment, and post-assignment survey, as well as the assigned system.

In the study, two groups of participants were asked to do an extra credit assignment (i.e., writing a summary of a given chapter) using two different systems: the proposed system (with feedback) and the baseline system (without feedback). Also, they were invited to participate in two online surveys before and after the assignment activity. The questionnaire contained both multiple-choice questions using a five-point Likert scale and open-ended questions. Both surveys were administered using SurveyMonkey, an online survey hosting site, and confidentiality was maintained using unique identifiers for respondents.

The valid participants were selected from the students who participated in all stages of this study, including responding to the pre-assignment survey, writing assignment using the designated system, and responding to the post-assignment survey. Two survey submissions were disqualified due to incomplete submissions, resulting in a total of 30 usable responses (15 for the experimental group; 15 for the control group) in data analysis. Due to the small number of complete responses, no meaningful inferential statistical tests could be performed, and statistical results are reported here for informational purposes only. Basic demographic information for the student participants is shown in Table 4.3.

Table 4.3 Participant Demographics for Pilot Study (N = 30)

Gender	17 Male (56.7%)	13 Female (43.3%)
Level of Study	2 Undergraduate (6.7%)	28 Graduate (93.3%)
Academic Program	13 Information Systems (43.3%)	
	9 Computer Science (30.0%)	
	3 Information Technology (10.0%)	
	3 Management (10.0%)	
	2 Electrical and Computer Engineering (6.7%)	
International Student	24 Yes (80.0%)	6 No (20.0%)
Students Reporting English as First Language	13 Yes (43.3%)	17 No (56.7%)

4.3.3 Results

Results of the study suggested that most students recognized the value of automated feedback in improving their learning. About half (46.7%) of students had previously never received automated feedback during the past year, though most of them recognized the importance of feedback for their learning and prefer real-time feedback on their coursework. The results further strengthened the need of an automated formative feedback system to support their learning and improve their coursework during the process of writing an assignment.

Our pilot study sought to understand student perceptions of the proposed prototype system. As analyzed, the prototype system was perceived to be a useful tool for supporting student learning of conceptual knowledge in the process of writing assignments. All the participants in the experimental group (100%) expressed that they enjoyed the writing and learning process using the automated feedback system, and also agreed that the writing assignment helped them learn some important concepts in this course. Most of them (96.7%) believed that they learned to see relationships between course concepts through the writing assignment. Most participants (96.7%) agreed or strongly agreed that automated feedback provided in the proposed system was helpful for improving their coursework and motivating their learning.

Overall, there were no statistically significant differences in perceived ease of use and perceived usefulness between the experimental and control group students in the study. Regarding perceived ease of use, the results were not surprising, because the design of system interface was largely identical, except for the request feedback button and a sidebar for displaying feedback in the proposed system. Regarding perceived usefulness, in the pilot study, the participants were only asked to use either the baseline or the experimental system for a single assignment; therefore, the results do not mean that the participants thought the systems were equally useful. This suggested to us that, to better compare the perceived usefulness, a repeated measures design for the main study would be a superior choice.

Although students' perceived ease of use or perceived usefulness were not significantly different between the experimental and control groups, none of the experimental group students commented negatively on the use of the proposed system to provide feedback on their writing assignments, and they expressed the positive contribution of the system to improve their learning and coursework. For example, below are some typical comments from the participants, "*easy to use, easy to understand, user-friendly, interactive, great interface, and fun system to learn the concepts with.*" Other participants said, "*It is really helpful in understanding the course better*"; "*Getting feedback side-by-side helped in improving my work*"; and "*Best way of learning, understanding the concepts, fun to learn…*" Overall, the experimental group students responded positively to using the feedback generation feature in the system.

The pilot study results implied that the prototype system had no critical usability issues and most of the experimental group students perceived the system positively and would recommend the system to their friends or classmates. Although the number of participants in this pilot study was too small to achieve statistical significance, the study results were helpful in evaluating the measurement instruments, and the use of the prototype system for the larger field study for the dissertation. Based on the survey results, the assignment was perceived as not difficult by all the student participants, and students who completed the assignment obtained the same amount of extra credit. Thus, to better understand how students use the provided feedback, it is necessary to use required or more difficult writing assignments in the main study. By doing so, students are expected to have more interactions with the system, so that the effects of the automated feedback on student learning can be further investigated.

4.3.4 Discussion

Based on the findings from the pilot study, we made the following improvements and modifications for the system and study instruments to be used in the main study: First, the major revision we made for the main study was the experimental design. To conduct further investigations of our research questions, the repeated measures experimental design was used in the main study across two semesters. The within-subjects design can allow students to directly compare the capabilities of the proposed system (with feedback) and the baseline system (without feedback) for the same type of assignment. Second, the survey instruments (one pre-assignment survey and two post-assignment surveys) were refined and several new questions were added into the second post-assignment survey for the repeated measures user study, which will be introduced further in the next chapter. Third, in the repeated measures user study, student participants were required to use two versions of systems to complete two writing assignments. In order not to overwhelm student participants with extra course work and affect the existing course syllabus prepared by the instructors, the extra credit assignment was replaced with the required writing assignments in the main study. The completion of these graded assignments was mandatory as part of the regular coursework. Finally, the system generated a few generic terms that carried little domain-specific information and were not considered as key concepts by the evaluators. We improved our system by extending additional stop words removal to exclude more generic terms and refined the system interface as necessary before the main study.

4.4 Summary

This chapter presented the results of the preliminary studies, which included the evaluation of extracted domain concepts, the protocol analysis to examine the system functions and usability, as well as the pilot study to test the instruments and research design for the main study. Information collected from the preliminary studies in this chapter was used to finalize the design details of the main study, which will be discussed in the next chapter.

CHAPTER 5

MAIN STUDY

5.1 Overview

The overarching objective of our research is to support student learning of conceptual knowledge in WTL activities by providing individual students with automated formative feedback generated by the Write-and-Learn prototype system. To evaluate whether the research objective has been accomplished successfully, the main study was conducted to investigate the following research questions further: (1) how effective is the automated formative feedback generated with the help of concept maps constructed from lecture slides and writing assignments; (2) whether and how students utilize the automated formative feedback in WTL activities; (3) to what extent the automated formative feedback on writing assignments can affect student learning outcomes; and (4) do students recommend future use of the proposed automated feedback system? Why or why not? For the first research question, the evaluation of extracted domain concepts has been discussed in Chapter 4. Study results based on the survey analysis and log analysis are discussed in this chapter.

5.2 Method

In this research, courses offered every semester with multiple sections were given higher priority so that the required and graded writing assignments could be repeated over the course of two semesters (Fall 2018 and Spring 2019). Two instructors teaching the same course (i.e., IS677 Information Systems Principles) offered by the Informatics Department at NJIT were willing to allow a repeated measures study in their classes across two semesters: one instructor teaching two sections in Fall 2018 and one section in Spring 2019;

another instructor teaching one section in Fall 2018 and one section in Spring 2019. Once the instructors and courses had been identified, we worked with the instructors to determine the time period during which research activities (e.g., pre-assignment survey, writing assignments, and post-assignment surveys) could be conducted.

In our main study, student participants were required to complete two writing assignments, and they were informed that the assignments as part of the regular coursework would be graded. The instructors selected three writing assignments from different chapters of the textbook for us to incorporate into the experimental design. To obtain richer repeated measures data to answer our research questions, one of the two assignments selected in each semester was different. Scheduling was approved by the instructors to ensure that our research was introduced and completed without disruption to the existing course syllabus. The writing assignments were in the form of online discussions that consisted of two parts: as the first part of the assignment, students needed to apply the course content to analyze an article related to a course topic, and then write a summary to address some given questions. Students were then required to post their summary in the discussion forum as the initial comments, so that other students could comment on each other for further discussion of the course topic. For our research purposes, we only used the first part of the online discussion (i.e., summary) as our research assignments that students needed to complete using the assigned systems. The instructions for these writing assignments are shown in Appendix B.

The main study asked students to complete three surveys: a pre-assignment survey prior to working on the first assignment and two post-assignment surveys after completing each of the two assignments. Although the completion of the graded assignment was mandatory, their participation in the surveys was entirely voluntary and was awarded extra credit. Students who did not wish to participate in the study were offered an equivalent alternative assignment to obtain the same amount of extra credit. The above information was explained in the consent form approved by the Institutional Review Board (IRB) of the university (IRB Protocol Number F366-18). A copy of the consent form and FERPA waiver can be found in Appendix A. After signing the consent forms, the participants were given basic instructions describing the research tasks, including pre-assignment survey, writing assignments, and post-assignment surveys, as well as the assigned systems.

As mentioned in Chapter 4, the survey instrument was designed based on constructs validated in prior research (Davis, 1989; Venkatesh et al., 2003) and adapted to the context of this research. This study incorporated a total of forty measurement items on the questionnaires to measure nine variables, which included six items for self-efficacy, two items for effort expectancy, five items for facilitating conditions, four items for perceived feedback effectiveness, four items for perceived enjoyment, five items for perceived ease of use, four items for perceived usefulness, nine items for perceived learning, and one item for recommendation for use. Additionally, two measures were included in the final post-assignment surveys after students used two versions of systems to do the required assignments. These measures captured students' opinions regarding perceived value of the generated feedback and system preference. All questionnaire items were measured using a five-point Likert scale. Appendix D lists the definition and the measurement items for each variable. All surveys in their entirety for this study can be found in the Appendix C.

In the main study, learning outcomes are measured in two dimensions: perceived learning and actual learning. "Perceived Learning (PL)" is defined as the degree to which a person believes that using a particular system improves his or her learning, which was operationalized through questions in the post-assignment surveys to capture student perceptions of their learning. “Actual Learning (AL)” refers to students’ assignment performance, which can be captured by instructors’ assignment grades (Suskie, 2018), for students who consented to release their assignment grades by signing the Family Educational Rights and Privacy Act (FERPA) waiver included in the consent form. In addition to assignment grades, it is suggested that evidence of student learning can be captured in different ways such as observations of student behavior, feedback from computer simulated tasks (e.g., information on patterns of actions), as well as student reflections on their values, attitudes and beliefs (Suskie, 2018). The focus of our study is to explore how the automated formative feedback can be provided to support student learning of conceptual knowledge in WTL activities. Previous studies (Boud & Molloy, 2013; Sadler, 1989) suggest that the only way to determine if learning results from feedback is for students to utilize the feedback in the assignment writing process to complete the feedback loop, i.e., to ensure that information is received and acted on. Thus, learning outcomes in our study are measured in two dimensions: actual learning is evaluated through log analysis of students’ actual system usage and assignment performance, and perceived learning is examined through survey analysis of students’ self-reported responses.

Students in each course section were randomly assigned into two groups prior to the study. Students in odd numbered groups used the proposed system for the first assignment; students in even numbered groups used the baseline system for the first assignment. The participants were registered with the systems at the beginning of each semester and provided with a username and a password to log on to the systems. Students willing to participate in the study also completed a pre-assignment survey and a post-assignment survey for the first assignment. Several weeks later, these same groups were instructed to use the system they did not already use for the second assignment. In other words, students in odd numbered groups used the baseline system for the second assignment; students in even numbered groups used the proposed system for the second assignment. A counterbalanced order was used to ensure that every student was randomly exposed to the system with or without feedback at least once. Finally, the participants completed the second post-assignment survey for the second assignment. To ensure that students did not complete the post-assignment surveys before doing the assignments, the survey links were not made available until after the due dates for the assignment submissions.

All the surveys were administered online using SurveyMonkey, and confidentiality was maintained using unique identifiers for respondents. The pre-assignment survey captured the participants' demographic data including gender, age, educational level, academic program, English proficiency and whether they were international students. Also, it captured the participants' perceptions, experiences, and frequencies of receiving feedback. The post-assignment surveys included several statements exploring student perceptions of the provided automated feedback and their satisfaction with the assigned systems. All questionnaire items were measured on a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5). Additionally, several new questions exploring system preference were added to the second post-assignment survey after students had completed two assignments using two different systems as displayed in Appendix C. In addition to the structured questions, the questionnaires included several open-ended questions on the perceived usability and general impressions of the proposed system, such as what aspects of the proposed system the students liked best and least.

From the main study, two types of data were collected: the subjective ratings provided by the student participants in the pre-assignment and post-assignment survey questionnaires, as well as activity data that the system generated and collected during the study, including log data, assignment drafts and finished assignments submitted by the student participants. All the data collected was kept confidential and no identifying information was released. Data from the pre- and post- assignment surveys was first analyzed individually and then was merged using the students' identifiers to correlate their responses in different surveys. Quantitative statistical methods were used to analyze the empirical data. We assessed missing values, outliers, and normality to ensure quality data for analysis: all usable responses were complete, and there were no extreme value outliers because of the ordinal data. Survey data was analyzed using IBM SPSS Version 24.0.0.0.

5.3 Participants

The main study was conducted in an introductory graduate course entitled "Information Systems Principles" over two semesters consecutively at NJIT. During the fall semester of 2018, a total of 65 students enrolling in three different sections of the course were recruited to take part in the study. Among them, 5 students could not complete the whole study and were therefore removed, resulting in a total of 60 participants. During the spring semester of 2019, a total of 40 students enrolling in two different sections of the course were recruited to participate in the study. Among them, 12 students could not complete the whole study and were therefore removed, resulting in a total of 28 participants. The total number of students who participated in the main study across two semesters is 88 (N=88). Table 5.1 shows the number of participants for each of all the five course sections over two semesters.

Table 5.1 Number of Participants by Course Section

Semester	Section	Number of Students	Number of Students who Participated in the Study	Number of Students who Completed the Study
Fall 2018	Section 1 (IS677-001)	24	23	22
	Section 2 (IS677-101)	22	21	21
	Section 3 (IS677-103)	28	21	17
	SUM	74	65	60
Spring 2019	Section 1 (IS677-002)	28	24	19
	Section 2 (IS677-104)	24	16	9
	SUM	52	40	28
TOTAL		126	105	88

Table 5.2 presents demographic information for the participants (N=88). Most participants (88.7%) ranged from 21 to 30 years old. Almost all of them were graduate students and only one of them was an undergraduate student, as they were recruited from a graduate-level course. Gender distribution was not equally distributed (about 64.8% male and 35.2% female); however, it was expected as the distribution was representative of the overall gender make-up of the university’s student population. Most of the participants (76.1%) self-identified as international students, and a little less than half of them (47.7%) stated that English was their first language.

Table 5.2 Participant Demographics for Main Study (N = 88)

	Characteristics	Percentage
Gender	Male	64.8%
	Female	35.2
Age	18-20	0.0
	21-24	48.9
	25-29	39.8
	30-39	11.4
	40-49	0.0
	>50	0.0
Level of Study	Undergraduate	1.1
	Graduate	98.9
Academic Program	Information Systems	45.5
	Computer Science	31.8
	Information Technology	9.1
	Management	8.0
	Others	5.7
Are You International Student?	Yes	76.1
	No	23.9
Is English Your First Language?	Yes	47.7
	No	52.3

5.4 Participants' Prior Experiences with Feedback and Self-Efficacy

In addition to the demographic information, several generic questions were asked to explore students' perceptions and previous experiences of receiving feedback in general, and how often they received feedback in different courses during the past year. As shown in Table 5.3, it is evident that the majority of students value the importance of feedback, since the combined percentage of "agree" and "strongly agree" exceeds the other opinions.

Table 5.3 Perceptions of Receiving Feedback (N = 88)

Questionnaire Items	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
Feedback is important to me.	0.0	1.1	2.3	39.8	56.8
Feedback can guide me to improve my coursework.	0.0	1.1	3.4	34.1	61.4
Feedback motivates me to study.	0.0	3.4	6.8	44.3	45.5
Feedback tells me what I need to do to improve my performance in a subject.	0.0	0.0	3.4	40.9	55.7
It is more important for me to know how to improve my coursework before I receive a particular grade.	0.0	2.3	5.7	35.2	56.8
I always read the feedback on my assignments.	1.1	0.0	4.5	31.8	62.5
I always use the feedback to improve my assignments.	0.0	0.0	9.1	38.6	52.3

(Note: SA=strongly agree; A=agree; N=neutral; D=disagree; SD=strongly disagree)

Table 5.4 shows the prior experiences that students had in receiving feedback from their instructors. Most students stated that their instructors were always willing to provide feedback (70.5%) and they provided enough information to make feedback useful (70.4%). Most of them said that the feedback they received was relevant to their goals as a student (81.9%) and related to the purpose of the assignment (84.1%). But only more than half of students (59.1%) expressed that they received enough feedback from their instructors.

Table 5.4 Experiences of Receiving Feedback (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I receive enough feedback from my instructors.	1.1	13.6	26.1	38.6	20.5
My instructors are always willing to provide feedback.	0.0	5.7	23.9	37.5	33.0
My instructors provide enough information to make feedback useful.	1.1	9.1	19.3	44.3	26.1
The feedback I receive is relevant to my goals as a student.	0.0	2.3	15.9	45.5	36.4
The feedback I receive is related to the purpose of the assignment.	0.0	3.4	12.5	46.6	37.5

Table 5.5 indicates the frequency percentages of receiving feedback in courses during the past year. About half of students reported that the frequency of receiving feedback from the instructors is less than once a month, including individual verbal feedback (60.2%), individual written comments (55.7%), group verbal feedback (44.3%), and group written comments (51.1%). It is interesting to note that 31.8% of student participants have never received automated feedback in all the courses they have taken during the past year, although most of them valued the importance of feedback.

Table 5.5 Frequency of Receiving Feedback in the Past Year (N = 88)

Questionnaire Item	Several times a day	Several times a week	Once a week	Several times a month	Once a month	Less than once a month	Never
	%	%	%	%	%	%	%
Individual verbal feedback from the instructor	0.0	4.5	19.3	15.9	22.7	20.5	17.0
Individual written comments from the instructor	0.0	2.3	18.2	23.9	25.0	19.3	11.4
Group verbal feedback from the instructor	0.0	6.8	22.7	26.1	22.7	11.4	10.2
Group written comments from the instructor	0.0	2.3	22.7	23.9	27.3	13.6	10.2
Peer feedback	0.0	17.0	17.0	28.4	11.4	17.0	9.1
Self-assessment	2.3	33.0	15.9	18.2	10.2	10.2	10.2
Automated feedback	2.3	6.8	14.8	10.2	14.8	19.3	31.8

In addition to students' perception of feedback in general, participants were asked several statements exploring their perceived self-efficacy and effort expectancy prior to working on the first assignment, which can help us understand participants' expectations of their performance on the assignment through the pre-assignment survey.

Self-Efficacy

In the study, student participants were required to use two versions of the systems to write two graded assignments in English. As most of them (76.1%) were international students, and about half of them (47.7%) stated that English was their first language, the first two items in the self-efficacy scale captured student perceptions regarding their capability to write the assignments in English after they read the assignment instructions. In addition to writing self-efficacy, the last two items in the scale captured their system-specific self-efficacy after they were given a system demo prior to the study. Table 5.6 shows that most of students felt confident writing the assignments in English and operating the system functions.

Table 5.6 Self-Efficacy (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I feel confident writing assignments in English.	0.0	0.0	6.8	37.5	55.7
I have the necessary skills for writing assignments in English.	0.0	1.1	4.5	37.5	56.8
I feel confident finding information in the system.	1.1	0.0	5.7	43.2	50.0
I have the necessary skills for using the system.	0.0	0.0	3.4	44.3	52.3
I feel confident using system features.	2.3	0.0	6.8	37.5	53.4
I feel confident operating system functions.	1.1	2.3	5.7	44.3	46.6

Effort Expectancy

Table 5.7 displays student perceptions of effort expectancy. About 45.4% of student participants felt the course was easy or very easy for them, and about 40.9% of them felt the writing assignments used in this study were easy or very easy. A little less than half of students found they were neither easy nor difficult.

Table 5.7 Effort Expectancy (N = 88)

Questionnaire Item	VE (1)	E (2)	N (3)	D (4)	VD (5)
	%	%	%	%	%
How easy/difficult do you find this course?	6.8	38.6	47.7	4.5	2.3
How easy/difficult do you find this writing assignment?	10.2	30.7	45.5	11.4	2.3

(Note: VE=very easy; E=easy; N=neither easy nor difficult; D=difficult; VD=very difficult)

Analysis of students' responses suggests that participants generally had positive perceptions about their ability to do well in the assignments after they reviewed the assignment instructions and system instructions.

5.5 Results

In our main study, student participants were required to complete two graded writing assignments. As part of the research activities, students who consented to participate in the study were asked to complete three surveys: a pre-assignment survey prior to working on the first assignment and two post-assignment surveys after completing each of the two assignments. This section presents the major findings from the survey analysis and log analysis to address our research questions.

5.5.1 Survey Results

In the study, participants were asked to complete two post-assignment surveys after completing each of the two assignments. Several variables from the first post-assignment survey were repeated in the second post-assignment survey administered after submission of the second assignment, which included facilitating conditions, perceived ease of use, perceived usefulness, perceived learning, and recommendation for use. In addition, to understand student perceptions of the generated feedback, measures related to feedback evaluation were included in the post-assignment surveys. Several additional questions were included in the second post-assignment survey, asking students to compare two systems and assess the extent to which they found the feedback to be of value. Additionally, several open-ended questions were added in the post-assignment surveys, asking what they liked least and most about each system.

5.5.1.1 Feedback Evaluation. The key feature of the proposed system is to provide individual students with real-time, actionable, and individualized feedback based on what they write during a WTL activity. As the first step to evaluate feedback effectiveness, the evaluation of extracted domain concepts was conducted with the help of two independent domain experts, which has been discussed in Chapter 4. Because the automated feedback was generated specifically for a particular version of student assignment draft and there were too many for experts to review, we relied on the survey to inform us the quality of the generated feedback. After the participants have used the proposed system to write the required assignment, they were asked to complete the post-assignment survey to provide their perceptions and attitudes towards the generated feedback in the system. The survey results are discussed below.

Perceived Feedback Effectiveness

Table 5.8 shows descriptive analysis of student perceptions regarding the effectiveness of automated feedback provided in the system. It shows that the majority of students agreed or strongly agreed that the feedback they received in the proposed system was relevant (78.4%), easy to understand (82.9%), accurate (76.1%) and useful (79.5%), which are the main measures used to evaluate the effectiveness of feedback in related work (Roscoe et al., 2018; Thurlings et al., 2013).

Table 5.8 Perceived Feedback Effectiveness (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
The content of feedback is relevant.	0.0	2.3	19.3	53.4	25.0
The content of feedback is easy to understand.	0.0	1.1	15.9	54.5	28.4
The content of feedback is accurate.	0.0	1.1	22.7	50.0	26.1
The content of feedback is useful.	0.0	1.1	19.3	54.5	25.0

Perceived Value of Feedback

To assess the extent to which the students found the automated feedback to be of value, several questions were added into the second post-assignment survey after they had finished the two assignments using the assigned systems. These scale items were negatively or positively worded to guard against common method bias. As shown in Table 5.9, only about 20% of students agreed or strongly agreed that they found the feedback to be a waste of time, irrelevant to their writing assignments, and useless to their learning, and hard to understand. About 71.6% of students agreed or strongly agreed that they found the feedback to be valuable to their learning.

Table 5.9 Perceived Value of Feedback (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I found the feedback to be a waste of time.	19.3	38.6	23.9	12.5	5.7
I found the feedback to be irrelevant to my writing assignment.	17.0	37.5	23.9	14.8	6.8
I found the feedback to be useless to my learning.	26.1	31.8	23.9	11.4	6.8
I found the feedback to be hard to understand.	20.5	36.4	23.9	12.5	6.8
I found the feedback to be valuable to my learning.	3.4	5.7	19.3	45.5	26.1
I preferred the No Feedback version of the system.	19.3	35.2	27.3	12.5	5.7

Perceived Enjoyment

The descriptive statistics in Table 5.10 show that most students enjoyed receiving feedback while writing the assignment (80.7%) and enjoyed the writing and learning process using the proposed automated feedback system (75%). Moreover, most students believed that they were motivated to do their best work (81.9%) and found using the system to be enjoyable (72.7%).

Table 5.10 Perceived Enjoyment (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I enjoy receiving the feedback while I am writing the assignment.	2.3	3.4	13.6	48.9	31.8
I am motivated to do my best work.	2.3	4.5	11.4	48.9	33.0
I enjoy the writing and learning process in the system.	3.4	2.3	19.3	39.8	35.2
I find using the system to be enjoyable.	4.5	2.3	20.5	43.2	29.5

Perceived Learning from Feedback

When asked to what extent the automated feedback helped students learn, most students gave positive responses as shown in Table 5.11: most participants agreed or strongly agreed that the provided automated feedback can motivate them to revise their assignment (77.3%), gain a better understanding of course concepts (83%), and improve their assignment (78.4%).

Table 5.11 Perceived Learning from Feedback (N = 88)

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I learned from the feedback during the process of writing the assignment.	0.0	1.1	19.3	55.7	23.9
The feedback motivated me to revise my assignment.	0.0	1.1	21.6	50.0	27.3
The feedback encouraged me to reflect on how I can improve my assignment.	0.0	1.1	21.6	51.1	26.1
The feedback helped me gain a better understanding of course concepts.	0.0	2.3	14.8	55.7	27.3
The feedback helped me find out what I do not know.	0.0	2.3	20.5	53.4	23.9
The feedback helped me improve the quality of my assignment.	0.0	2.3	19.3	50.0	28.4

Perceived System Usage

In the proposed system, students can request feedback whenever they need some hints or suggestions. Based on the analysis of survey data, all the students self-reported that they have requested the feedback in the course of writing the assignments as listed in Table 5.12. In addition to requesting the feedback, it is also important to understand if the students read them or not and how carefully they reviewed the provided feedback. Only one student reported he or she did not read the feedback. When asked how carefully they read the feedback provided in the system, the majority of students across different course sections reported that they read the feedback carefully when they did the assignment.

Table 5.12 Descriptive Statistics of Perceived System Usage (N=88)

<table>
<tr><td>SU1</td><td colspan="7">Did you request feedback from the Write-and-Learn system while writing your assignment?</td></tr>
<tr><td></td><td colspan="3">Choice</td><td colspan="2">Yes</td><td colspan="2">No</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Fall 2018</td><td colspan="2">Assignment 1 (N=30)</td><td colspan="2">30 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td colspan="2">Assignment 2 (N=30)</td><td colspan="2">30 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Spring 2019</td><td colspan="2">Assignment 1 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td colspan="2">Assignment 2 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td>SU2</td><td colspan="7">Did you read the feedback provided in the Write-and-Learn system?</td></tr>
<tr><td></td><td colspan="3">Choice</td><td colspan="2">Yes</td><td colspan="2">No</td></tr>
<tr><td></td><td rowspan="2">Fall 2018</td><td colspan="2">Assignment 1 (N=30)</td><td colspan="2">29 (96.7%)</td><td colspan="2">1 (3.3%)</td></tr>
<tr><td></td><td colspan="2">Assignment 2 (N=30)</td><td colspan="2">30 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td></td><td rowspan="2">Spring 2019</td><td colspan="2">Assignment 1 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td></td><td colspan="2">Assignment 2 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td>SU3</td><td colspan="7">How carefully did you read the feedback provided in the Write-and-Learn system?</td></tr>
<tr><td></td><td colspan="2">Choice</td><td>Not at all carefully</td><td>Not so carefully</td><td>Somewhat carefully</td><td>Very carefully</td><td>Extremely carefully</td></tr>
<tr><td></td><td rowspan="2">Fall 2018</td><td>Assignment 1 (N=30)</td><td>1 (3.3%)</td><td>1 (3.3%)</td><td>10 (33.3%)</td><td>11 (36.7%)</td><td>7 (23.3%)</td></tr>
<tr><td></td><td>Assignment 2 (N=30)</td><td>0 (0.0%)</td><td>4 (13.3%)</td><td>7 (23.3%)</td><td>14 (46.7%)</td><td>5 (16.7%)</td></tr>
<tr><td></td><td rowspan="2">Spring 2019</td><td>Assignment 1 (N=14)</td><td>1 (7.1%)</td><td>1 (7.1%)</td><td>4 (28.6%)</td><td>7 (50.0%)</td><td>1 (7.1%)</td></tr>
<tr><td></td><td>Assignment 2 (N=14)</td><td>1 (7.1%)</td><td>2 (14.3%)</td><td>5 (35.7%)</td><td>3 (21.4%)</td><td>3 (21.4%)</td></tr>
<tr><td>SU4</td><td colspan="7">Did you revise your assignment based on the provided feedback?</td></tr>
<tr><td></td><td colspan="3">Choice</td><td colspan="2">Yes</td><td colspan="2">No</td></tr>
<tr><td></td><td rowspan="2">Fall 2018</td><td colspan="2">Assignment 1 (N=30)</td><td colspan="2">29 (96.7%)</td><td colspan="2">1 (3.3%)</td></tr>
<tr><td></td><td colspan="2">Assignment 2 (N=30)</td><td colspan="2">29 (96.7%)</td><td colspan="2">1 (3.3%)</td></tr>
<tr><td></td><td rowspan="2">Spring 2019</td><td colspan="2">Assignment 1 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
<tr><td></td><td colspan="2">Assignment 2 (N=14)</td><td colspan="2">14 (100%)</td><td colspan="2">0 (0.0%)</td></tr>
</table>

5.4.1.2 System Evaluation. In the study, several variables from the first post-assignment survey were repeated in the second post-assignment survey after submission of the second assignment, which include facilitating conditions, perceived ease of use, perceived usefulness, perceived learning, and recommendation for use. The means and distributions of these measures are reported in this section. Also, to investigate whether there were any interactions between students' experiences during the two assignments and their assigned systems, a two-way repeated measures ANOVA was conducted for each repeated variable. Facilitating conditions, perceived ease of use, and perceived learning revealed no interaction effects between the conditions of assigned system and assignment. Perceived usefulness and recommendation for use showed an interaction effect with the students' assigned system at the 90% confidence level. The interactions of perceived usefulness and recommendation of use and the order of use of assigned system was not significant. Each of these results is examined in more detail as below.

Univariate Analysis of Facilitating Conditions

As mentioned earlier, students were given basic instructions about the writing assignments and the assigned systems prior to the study. Table 5.13 shows the facilitating conditions of using the assigned systems to write the assignments in this study. None of the five items displayed statistically significant differences between the two systems.

Table 5.13 Facilitating Conditions

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
The grading rubric is fair and explicit enough.					
Baseline System (N=88) Mean = 3.94, SD = 0.835	1.1	4.5	17.0	53.4	23.9
Proposed System (N=88) Mean = 3.89, SD = 0.794	1.1	2.3	23.9	52.3	20.5
The assignment instruction is clear enough.					
Baseline System (N=88) Mean = 4.23, SD = 0.690	0.0	1.1	11.4	51.1	36.4
Proposed System (N=88) Mean = 4.22, SD = 0.718	0.0	3.4	6.8	54.5	35.2
When I need help to use the system, guidance is available to me.					
Baseline System (N=88) Mean = 4.13, SD = 0.740	0.0	5.7	4.5	61.4	28.4
Proposed System (N=88) Mean = 4.17, SD = 0.746	0.0	4.5	6.8	55.7	33.0
Special instruction concerning the system use (e.g., system demo, system instructions) is available to me.					
Baseline System (N=88) Mean = 4.17, SD = 0.761	0.0	3.4	11.4	50.0	35.2
Proposed System (N=88) Mean = 4.23, SD = 0.754	0.0	2.3	12.5	45.5	39.8
A specific person/group is available for assistance with any difficulties related with the system use.					
Baseline System (N=88) Mean = 3.95, SD = 0.883	0.0	6.8	20.5	43.2	29.5
Proposed System (N=88) Mean = 4.10, SD = 0.788	0.0	3.4	15.9	47.7	33.0
$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Interaction of Facilitating Conditions and Assigned System

A two-way repeated measures ANOVA was conducted to explore the interaction of facilitating conditions from two assignments and the systems to which they were assigned. The interaction of facilitating conditions and assigned system was not significant, with $F(1) = 1.043$, $p = 0.309$. A further exploration of the means of the four conditions is shown in Table 5.14. Regardless of assignment, the means of the facilitating conditions for students using the proposed system were higher than the means for students using the baseline system.

Table 5.14 Means of Facilitating Conditions Repeated Measures

Facilitating Conditions	Baseline System	Proposed System	Both Systems
Assignment 1	20.32 SD = 3.094	20.61 SD = 3.377	20.47 SD = 3.223
Assignment 2	20.20 SD = 3.414	20.50 SD = 3.092	20.35 SD = 3.241
Both Assignments	20.26 SD = 3.239	20.56 SD = 3.219	

Univariate Analysis of Perceived Ease of Use

To better understand the students' perception of the usability of the proposed system, this study collected the participants' responses regarding "perceived ease of use" and "perceived usefulness". As for the perceived ease of use, it is found that most students gave positive feedback concerning the proposed system, although none of the five items displayed statistically significant differences between the two systems. As shown in Table 5.15, about 80% of students agreed or strongly agreed that the proposed automated feedback system was easy to use, user-friendly, stable, and easy to understand.

Table 5.15 Perceived Ease of Use

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
The system is easy to use.					
Baseline System (N=88) Mean = 4.01, SD = 0.953	3.4	4.5	10.2	51.1	30.7
Proposed System (N=88) Mean = 4.15, SD = 0.824	0.0	5.7	10.2	47.7	36.4
The system is user-friendly.					
Baseline System (N=88) Mean = 3.98, SD = 1.039	3.4	6.8	13.6	40.9	35.2
Proposed System (N=88) Mean = 4.08, SD = 0.791	0.0	3.4	17.0	47.7	31.8
The operation of the system is stable.					
Baseline System (N=88) Mean = 3.97, SD = 1.066	4.5	6.8	10.2	44.3	34.1
Proposed System (N=88) Mean = 4.09, SD = 0.768	0.0	3.4	14.8	51.1	30.7
My interaction with the system is clear and understandable.					
Baseline System (N=88) Mean = 4.00, SD = 0.858	1.1	3.4	19.3	46.6	29.5
Proposed System (N=88) Mean = 4.05, SD = 0.843	0.0	6.8	12.5	50.0	30.7
Interacting with the system does not require a lot of my mental effort.					
Baseline System (N=88) Mean = 4.07, SD = 0.932	1.1	6.8	12.5	43.2	36.4
Proposed System (N=88) Mean = 4.14, SD = 0.776	0.0	3.4	13.6	48.9	34.1
$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Interaction of Perceived Ease of Use and Assigned System

A two-way repeated measures ANOVA was conducted to explore the interaction of perceived ease of use from two assignments and the systems to which they were assigned. The interaction of perceived ease of use and assigned system was not significant, with $F(1) = 1.699$, $p = 0.195$. A further exploration of the means of the four conditions is shown in Table 5.16, where the means of the perceived ease of use for students using the proposed system were higher than the means for students using the baseline system regardless of assignment.

Table 5.16 Means of Perceived Ease of Use Repeated Measures

Perceived Ease of Use	Baseline System	Proposed System	Both Systems
Assignment 1	20.09 SD = 3.640	20.82 SD = 3.391	20.45 SD = 3.516
Assignment 2	19.95 SD = 4.808	20.18 SD = 3.280	20.07 SD = 4.093
Both Assignments	20.02 SD = 4.240	20.50 SD = 3.332	

Univariate Analysis of Perceived Usefulness

In terms of the perceived usefulness, most students agreed or strongly agreed that the proposed system improved their learning performance (72.8%) and increased their academic productivity (71.6%). As listed in Table 5.17, most students agreed or strongly agreed that the system made it easier to study course concepts (80.7%) and was useful in learning course concepts (77.3%). To further explore the difference, each item in the perceived usefulness scale was analyzed individually and the last two items resulted in statistically significant differences between the proposed system and the baseline system, which will be discussed further in Chapter 6.

Table 5.17 Perceived Usefulness

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
Using the system would improve my learning performance.					
Baseline System (N=88) Mean = 3.73, SD = 0.991	3.4	3.4	28.4	39.8	22.7
Proposed System (N=88) Mean = 3.98, SD = 0.788	0.0	2.3	25.0	45.5	27.3
Using the system would increase my academic productivity.					
Baseline System (N=88) Mean = 3.70, SD = 1.019	3.4	8.0	26.1	39.8	22.7
Proposed System (N=88) Mean = 3.98, SD = 0.802	0.0	2.3	26.1	43.2	28.4
Using the system would make it easier to study course concepts.					
Baseline System (N=88) Mean = 3.72, SD = 1.028*	2.3	11.4	22.7	39.8	23.9
Proposed System (N=88) Mean = 4.09, SD = 0.753*	0.0	2.3	17.0	50.0	30.7
t = -3.304, p = 0.001					
I find the system to be useful in learning course concepts.					
Baseline System (N=88) Mean = 3.65, SD = 1.062*	3.4	12.5	21.6	40.9	21.6
Proposed System (N=88) Mean = 4.06, SD = 0.717*	0.0	0.0	22.7	48.9	28.4
t = -3.124, p = 0.002					
$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Interaction of Perceived Usefulness and Assigned System

A two-way repeated measures ANOVA was conducted to explore the interaction of perceived usefulness from two assignments and the systems to which they were assigned. The interaction of perceived usefulness and assigned system was significant at the 90% confidence level, with F (1) = 2.985, p = 0.087. The results suggest that there was an interaction between the system used and the assignment in terms of perceived usefulness. Students who used the proposed system responded more positively to the statement about making it easier to study course concepts and being useful in learning course concepts.

A further exploration of the means of the four conditions is shown in Table 5.18. As shown, the means of the perceived usefulness variable for students using the baseline system was lower (regardless of assignment) than the means for students using the proposed system. The interaction of perceived usefulness and the order of use of the assigned system was not significant, with F (1) = 0.280, $p = 0.598$.

Table 5.18 Means of Perceived Usefulness Repeated Measures

Perceived Usefulness	Baseline System	Proposed System	Both Systems
Assignment 1	15.50 SD = 4.032	15.93 SD = 2.937	15.72 SD = 3.513
Assignment 2	15.52 SD = 3.317	16.27 SD = 2.509	15.90 SD = 2.948
Both Assignments	15.51 SD = 3.670	16.10 SD = 2.721	

Univariate Analysis of Perceived Learning

In the post-assignment surveys, students were asked to provide their perceptions of learning from the writing assignments, after they have used the two versions of system to finish the two assignments. As shown in Table 5.19, the results revealed no statistically significant difference in perceived learning between the two systems.

Table 5.19 Perceived Learning

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
The writing assignment helped me learn some important concepts in this course.					
Baseline System (N=88) Mean = 4.02, SD = 0.816	1.1	3.4	14.8	53.4	27.3
Proposed System (N=88) Mean = 4.17, SD = 0.572	0.0	0.0	9.1	64.8	26.1
I learned to see relationships between course concepts through the writing assignment. (For example, “Six Sigma” is closely related to “Total Quality Management”.)					
Baseline System (N=88) Mean = 4.02, SD = 0.802	1.1	2.3	17.0	52.3	27.3
Proposed System (N=88) Mean = 4.13, SD = 0.708	0.0	1.1	15.9	52.3	30.7
I developed an improved ability to integrate course concepts through the writing assignment.					
Baseline System (N=88) Mean = 3.99, SD = 0.809	1.1	3.4	15.9	54.5	25.0
Proposed System (N=88) Mean = 4.07, SD = 0.740	0.0	2.3	17.0	52.3	28.4
$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Interaction of Perceived Learning and Assigned System

A two-way repeated measures ANOVA was conducted to explore the interaction of perceived learning from two assignments and the systems to which they were assigned. The interaction of perceived learning and assigned system was not significant, with $F(1) = 2.145$, $p = 0.146$. A further exploration of the means of the four conditions is shown in Table 5.20. As displayed, the means of the perceived learning for students using the proposed system were higher than the means for students using the baseline system regardless of assignment.

Table 5.20 Means of Perceived Learning Repeated Measures

Perceived Learning	Baseline System	Proposed System	Both Systems
Assignment 1	11.75 SD = 2.243	12.39 SD = 1.820	12.07 SD = 2.056
Assignment 2	12.32 SD = 2.143	12.34 SD = 1.804	12.33 SD = 1.969
Both Assignments	12.03 SD = 2.200	12.36 SD = 1.802	

One-Way ANOVA for Differences in Perceived Learning Between Instructors

As instructors play an important role in student learning, a one-way ANOVA was conducted to evaluate the differences in perceived learning when students used the proposed system to complete the writing assignments between the course sections taught by two instructors. Results show that there were no statistically significant differences in perceived learning between the course sections taught by two different instructors. Table 5.21 lists the perceived learning between courses taught by two different instructors, their mean differences, and significance.

Table 5.21 Differences in Perceived Learning Between Instructors

Variable	Instructor 1	Instructor 2	Mean Difference (1-2)
Perceived Learning F (87) = 0.150, p=0.700	Section 001, Section 101, 002	Section 103, Section 104	0.46
** p < 0.05, ** p < 0.01, *** p < 0.001, (ns) = not significant*			

Univariate Analysis of Recommendation of Use

The descriptive statistics as listed in Table 5.22 indicate that the majority (79.5%) of students would recommend doing this type of assignment with the proposed automated feedback system for future classes. However, there was no statistically significant difference between the two systems at the 95% confidence level.

Table 5.22 Recommendation of Use

Questionnaire Item	SD (1)	D (2)	N (3)	A (4)	SA (5)
	%	%	%	%	%
I would recommend doing the assignment with this system for future classes.					
Baseline System (N=88) Mean = 3.89, SD = 1.022	4.5	3.4	20.5	42.0	29.5
Proposed System (N=88) Mean = 4.09, SD = 0.705	0.0	0.0	20.5	50.0	29.5
$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Interaction of Recommendation of Use and Assigned System

A two-way repeated measures ANOVA was conducted to explore the interaction of recommendation of use from two assignments and the systems to which they were assigned. The interaction of recommendation of use and assigned system was significant at the 90% confidence level, with F (1) = 3.891, $p = 0.051$. As displayed in Table 5.23, the means of the recommendation of use for students using the proposed system were higher than the means for students using the baseline system regardless of assignment. The interaction of recommendation of use and the order of use of the assigned system was not significant, with F (1) = 0.190, $p = 0.664$.

Table 5.23 Means of Recommendation of Use Repeated Measures

Recommendation of Use	Baseline System	Proposed System	Both Systems
Assignment 1	3.82 SD = 1.063	4.11 SD = 0.722	3.97 SD = 0.915
Assignment 2	3.95 SD = 0.987	4.07 SD = 0.695	4.01 SD = 0.851
Both Assignments	3.89 SD = 1.022	4.09 SD = 0.705	

Univariate Analysis of System Preference

In the second post-assignment survey, students were asked to compare the system they used for the second assignment with the system they used for the first assignment, indicating if they liked the second system less or more. The question was included by the following statement: "*This was the second assignment of this type which you completed for this class. Compare your experience in the second time to the first time you did this type of assignment and respond to the following statements: I liked the second system…"* The question provided responses measured on a five-point Likert scale as follows: a great deal less (1), slightly less (2), about the same (3), slightly more (4), and a great deal more (5). To determine if there were any differences in preferences of the systems they used, an independent samples t-test was conducted. As displayed in Table 5.24, there was no statistically significant difference in the preferences for the two systems, although the proposed system has a higher mean than the baseline system as the preferred system.

Table 5.24 System Preference

I liked the second system …				
	Mean	SD	t statistic	Sig.
Both Systems	3.56	1.004	33.229	0.000
Second System Used Was …				
Baseline System	3.36	1.080	-1.829	0.091
Proposed System	3.75	0.892		
1=A great deal less, 2=Slightly less, 3=About the same, 4=Slightly more, 5=A great deal more				

As mentioned earlier, the writing assignments used in the main study were in the form of online discussions. For our research purposes, we only asked students to complete the first part of the online discussions (i.e., summary) using the assigned systems. After they completed the summary in the assigned systems, students were then required to post their summary as the initial comments in the discussion forum on Moodle, so that other students could comment on each other for further discussion of the course topic. Throughout the studies, the students were involved with three systems: the baseline system, the proposed system and Moodle. It is possible that students might be confused about which system we referred to in the statement, as we intentionally used "the second system" in the question we asked. Another possible explanation is that the interfaces of the proposed system and the baseline system were identical. Students had the option to request feedback at any time they needed. If students used the proposed system the same way as the baseline system, it might be difficult for them to rate their system preference, especially when the questions were asked after they have used both systems for a while.

To explore the research questions further, the second post-assignment survey included several open-ended questions asking students to explain what they liked best and least about using the system for this type of assignment. Students' reactions provided interpretation and deeper understanding of the results from the quantitative assessment of the proposed system. When asked what they liked best about using the proposed system for this type of writing assignment, students pointed out various reasons. In general, the students responded positively to using the feedback generation feature in the system. For example, one student stated, "*The feedback option for the article input is one of the best features in the system.*" Another student said, "*The concepts listed on the left of the screen, the PowerPoint being integrated in the online system and auto saving of the draft are three things I like about the system.*" One student commented that the system is "*quite useful in learning main concepts while doing the assignment and allows us to do the assignment to the best of our effort by displaying keywords.*" Another student mentioned, "*The* *integration with the current chapter, the recommendation of terms from the current chapter, and the simplicity.*"

A few students liked the fact that the feedback "encouraged to do better work" and "think outside-the-box, helped me use relevant keywords and concepts." Other comments included, "It showed me topics and words that I could have used. Acted as a suggestion." Few students went on to express their ideas as "It helped me to include the concepts that I missed to pay attention to." "It provided useful concepts from the chapter" and also "proper and clear feedback, improving my knowledge." Some other comments were related to the usability of the system, such as "system interface is very easy to understand and helps understanding the measures about the requirements of articles easily because of the prompt guidelines. It is very user friendly, easy to understand and an instructive system." Thus, the qualitative data obtained through the open-ended questions shows that the proposed system was viewed by most students as a helpful tool to support their learning and improve their coursework.

In response to what they liked least about using the proposed system for this type of assignment, students focused on the ease of connecting with their classmates. As the assignments used in the study were online discussions, one student mentioned it would be better if they could view the comments posted by other students directly, and another student wished that a chatroom should be integrated in the system. In response to what they liked least about using the baseline system for this type of assignment, one student commented, "*the first version of the system (the Write-and-Learn system) was more educational, and the second version (the baseline system) was just for posting assignment, so there is really no way it could help someone learn.*"

5.5.2 Log Analysis

When students used the assigned systems to write the assignments, the system inputs of each participant including the log data, assignment drafts and finished assignments can be automatically collected. To understand whether and how students utilized the automated feedback generated in the system, analysis of log data was conducted.

5.5.2.1 Actual System Usage. Although the students were encouraged to request feedback and most of them self-reported that they read the feedback generated by the system, it is more important to understand whether they actually revised their assignments after requesting the feedback. Analysis of log data shows that all participants requested feedback at least once when they used the proposed system, which is consistent with what they claimed in the survey. Based on the log analysis of students' actual usage of the system, during the fall semester of 2018, 29 out of 30 participants in Group 1 revised the assignment after requesting feedback for the first assignment, and 27 out of 30 in Group 2 made revisions after requesting feedback for the second assignment. During the spring semester of 2019, 13 out of 14 participants in Group 1 revised the assignment after requesting feedback for the first assignment, and 14 out of 14 in Group 2 made revisions after requesting feedback for the second assignment.

As mentioned earlier, the form of the assignment was online discussion. After using the assigned system to write the assignment, they still needed to post their initial comments on Moodle for other students' comments before the discussion forum was closed. Among the five students who did not revise their assignments after requesting feedback, three of them seemed to finish their assignments in a rush, because the time they spent on writing the assignment was less than 15 minutes, which was far below the average (about 35 minutes). Analysis of log data suggests that two of the three students did not start working on the assignment until it was almost due. In addition, one student did not make revisions after requesting the feedback, but several key concepts can be identified from the student's assignment submission. One possible explanation was that if the student felt comfortable about his or her assignment, there was no need to make further revisions and the student could submit the assignment once finished. Finally, the last student who did not revise the assignment after requesting the feedback, his or her ratings for the feedback effectiveness were neutral and no further comment was left to explain what aspects the student liked or disliked about the system.

In order to understand how students utilized the automated feedback when they wrote assignments, the system logs were utilized which reflected how students actually used the system. The recorded log files included: (1) *Number of Feedback Requests*: obtained by counting the total number of times a student requested feedback during the process of writing a given assignment using the proposed system; and (2) *Number of Revisions Made*: obtained by counting the total number of times a student revises his or her assignment ***after*** requesting feedback. Besides, the students' assignment drafts and submitted assignments were automatically collected via the system.

Table 5.25 displays the descriptive statistics of actual system usage across two semesters. During the fall semester of 2018, when the students (N=60) used the proposed system, the total number of attempts to request feedback in the course of writing the assignment was 439, with a mean of 7.32 and a standard deviation of 6.248; the total number of revisions made after requesting the feedback was 118, with a mean of 1.97 and a standard deviation of 1.605. During the spring semester of 2019, when the students

(N=28) used the proposed system, the total number of attempts to request feedback in the course of writing the assignment are 122, with a mean of 4.36 and a standard deviation of 2.725; the total number of revisions made after requesting the feedback was 48, with a mean of 1.71 and a standard deviation of 1.049. One possible explanation is that the durations of these online discussions were shortened due to the change of course schedule by the instructors in the spring semester of 2019. As a result, most participants spent less time on completing the assignments~~.~~

Table 5.25 Descriptive Statistics of Actual System Usage by Semester

Variables	Semester	Assignment	Total	Mean	SD	Minimum	Maximum
Attempts of Feedback Requested	Fall 2018	Assignment 1 (N=30)	250	8.33	6.804	1	27
		Assignment 2 (N=30)	189	6.30	5.566	1	22
		Assignment 1 and 2 (N=60)	439	7.32	6.248	1	27
	Spring 2019	Assignment 1 (N=14)	73	5.21	3.262	1	13
		Assignment 2 (N=14)	49	3.50	1.787	1	7
		Assignment 1 and 2 (N=28)	122	4.36	2.725	1	13
	Two Semesters	All Assignments (N=88)	561	6.38	5.547	1	27
Number of Revisions Made	Fall 2018	Assignment 1 (N=30)	63	2.10	1.689	0	7
		Assignment 2 (N=30)	55	1.83	1.533	0	6
		Assignment 1 and 2 (N=60)	118	1.97	1.605	0	7
	Spring 2019	Assignment 1 (N=14)	26	1.86	1.351	0	5
		Assignment 2 (N=14)	22	1.57	0.646	1	3
		Assignment 1 and 2 (N=28)	48	1.71	1.049	0	5
	Two Semesters	All Assignments (N=88)	166	1.89	1.450	0	7

As each participant was asked to complete one of the two writing assignments using the proposed system, the mean values of the above variables were compared for these two different assignments. As for the attempts of feedback during the fall semester of 2018, the mean value is 8.33 for Group 1 who did the first assignment with the proposed system, and 6.30 for Group 2 who did the second assignment with the proposed system respectively. As for the attempts of feedback requested during the spring semester of 2019, the mean value is 5.21 for Group 1 who did the first assignment with the proposed system, and 3.50 for Group 2 who did the second assignment with the proposed system respectively. It is found that the students made more requests for feedback during the process of writing the first assignment. Results suggest that when students used the proposed system, there were no statistically significant differences between these two assignments in terms of number of revisions made: 2.10 and 1.83 for the fall semester of 2018, and 1.86 and 1.57 for the spring semester of 2019, respectively.

Table 5.26 displays the descriptive statistics of actual system usage in terms of the number of attempts to request feedback and number of revisions made, based on the analysis of different course sections taught by two instructors across two semesters. As for the average number of times feedback was requested, there were significant differences between the course sections taught by two instructors across two semesters: 7.32 for instructor one and 4.12 for instructor two. There were no statistically significant differences in the average number of revisions made between the course sections taught by two different instructors.

Table 5.26 Descriptive Statistics of Actual System Usage by Instructor

Variables	Instructor	Assignment	Total	Mean	SD	Minimum	Maximum
Attempts of Feedback Requested	Instructor One	Fall 2018 (N=43)	358	8.33	6.732	1	27
		Spring 2019 (N=19)	96	5.05	2.896	1	13
		Two Semesters (N=62)	454	7.32	5.999	1	27
	Instructor Two	Fall 2018 (N=17)	81	4.76	3.914	1	13
		Spring 2019 (N=9)	26	2.89	1.833	1	5
		Two Semesters (N=26)	107	4.12	3.421	1	13
Number of Revisions Made	Instructor One	Fall 2018 (N=43)	77	1.79	1.473	0	7
		Spring 2019 (N=19)	36	1.89	1.197	0	5
		Two Semesters (N=62)	113	1.82	1.385	0	7
	Instructor Two	Fall 2018 (N=17)	41	2.41	1.873	0	6
		Spring 2019 (N=9)	12	1.33	0.500	1	2
		Two Semesters (N=26)	53	2.04	1.612	0	6

5.5.2.2 Assignment Performance. Learning outcomes are typically measured in two dimensions: perceived learning and actual learning. The focus of our study is to explore how the automated formative feedback can be provided to support student learning of conceptual knowledge in WTL activities. Previous studies (Boud & Molloy, 2013; Sadler, 1989) suggest that the only way to determine if learning results from feedback is for students to utilize the feedback in the assignment writing process to complete the feedback loop, i.e., to ensure that information is received and acted on. It is also pointed out that feedback cannot be judged as effective unless students can use the feedback to produce improved work (Boud, 2000). Thus, in addition to capturing students' perceived learning through survey responses, log analysis was conducted to examine students' assignment performance. To examine the effects of automated feedback on student learning, we calculated the number of key concepts and established concept relationships that can be identified in all the writing assignments completed using the proposed system and baseline system, respectively. Table 5.27 displays the descriptive statistics of assignment performance for each assignment across semesters. It is evident that students spent less time using the assigned system to complete the assignment during the spring semester of 2019, which might affect their assignment performance.

Table 5.27 Descriptive Statistics of Assignment Performance by System

Semester	Assignment	Variables	Total	Mean	SD	Minimum	Maximum
Fall 2018	Assignments Using the Baseline System (N=60)	Number of Concepts	173	2.88	1.354	1	7
		Number of Concept Relationships	1	0.02	0.129	0	1
	Assignments Using the Proposed System (N=60)	Number of Concepts	269	4.48	2.251	1	11
		Number of Concept Relationships	12	0.20	0.443	0	2
Spring 2019	Assignments Using the Baseline System (N=28)	Number of Concepts	54	1.93	1.245	0	5
		Number of Concept Relationships	2	0.07	0.262	0	1
	Assignments Using the Proposed System (N=28)	Number of Concepts	85	3.04	1.347	1	7
		Number of Concept Relationships	5	0.18	0.390	0	1
Two Semesters	Assignments Using the Baseline System (N=88)	Number of Concepts	227	2.58	1.387	0	7
		Number of Concept Relationships	3	0.03	0.183	0	1
	Assignments Using the Proposed System (N=88)	Number of Concepts	354	4.02	2.112	1	11
		Number of Concept Relationships	17	0.19	0.425	0	2

To evaluate the results in a consistent way, the lists of key concepts and established concept relationships generated by the proposed system were used to produce the results. As shown in Table 5.28, the differences between the mean values of the number of concepts included in the assignments completed using the proposed system and the baseline system were statistically significant. Besides, the differences between the number of concept relationships identified in the assignments completed using the proposed system, and the number of concept relationships identified in the assignments completed using the baseline system were also statistically significant. As there were the five students who did not revise their assignments after requesting feedback, the average of number of concepts and concept relationships included in the submitted assignments would be even higher if we only considered the 83 students who made the revisions of their assignments.

Table 5.28 Differences in Assignment Performance by System

Variables	System	Mean and SD	Results
Number of Concepts	Baseline	2.58, SD = 1.387	t = 8.056, df = 87***
	Proposed	4.02, SD = 2.112	
Number of Concept Relationships	Baseline	0.03, SD = 0.183	t = 3.751, df = 87***
	Proposed	0.19, SD = 0.425	
* p < 0.05, ** p < 0.01, *** p < 0.001, (ns) = not significant			

Table 5.29 displays the descriptive statistics of assignment performance, based on the analysis of different course sections taught by two instructors across two semesters.

Table 5.29 Descriptive Statistics of Assignment Performance by Instructor

Instructor	Assignment	Variables	Total	Mean	SD	Minimum	Maximum
Instructor One	Assignment Using the Baseline System (N=62)	Number of Concepts	165	2.66	1.305	1	6
		Number of Concept Relationships	3	0.05	0.216	0	1
	Assignment Using the Proposed System (N=62)	Number of Concepts	251	4.05	1.903	1	11
		Number of Concept Relationships	13	0.21	0.449	0	2
Instructor Two	Assignment Using the Baseline System (N=26)	Number of Concepts	62	2.38	1.577	0	7
		Number of Concept Relationships	0	0.00	0.000	0	0
	Assignment Using the Proposed System (N=26)	Number of Concepts	103	3.96	2.584	1	9
		Number of Concept Relationships	4	0.15	0.368	0	1

In Table 5.30, results show that the differences in assignment performance, in terms of the number of concepts and concept relationships included in the assignments completed using the proposed system and completed using the baseline system, were statistically significant.

Table 5.30 Differences in Assignment Performance by Instructor

Instructor	Variables	System	Mean and SD	Results
Instructor One	Number of Concepts	Baseline	2.66, SD = 1.305	t = 6.729, df = 61***
		Proposed	4.05, SD = 1.903	
	Number of Concept Relationships	Baseline	0.05, SD = 0.216	t = 3.078, df = 61**
		Proposed	0.21, SD = 0.449	
Instructor Two	Number of Concepts	Baseline	2.38, SD = 1.577	t = 4.378, df = 25***
		Proposed	3.96, SD = 2.584	
	Number of Concept Relationships	Baseline	0.00, SD = 0.000	t = 2.132, df = 25*
		Proposed	0.15, SD = 0.368	
* p < 0.05, ** p < 0.01, *** p < 0.001, (ns) = not significant				

To evaluate the differences in assignment performance when students used the proposed system to complete the writing assignments between the course sections taught by two instructors, a one-way ANOVA was conducted. In Table 5.31, results show that there were no statistically significant differences in assignment performance between the course sections taught by two different instructors.

Table 5.31 Differences in Assignment Performance Between Instructors

Variable	Instructor 1	Instructor 2	Mean Difference (1-2)
Number of Concepts F (87) = 0.031 *p*=ns	Section 001, Section 101, 002	Section 103, Section 104	0.09
Number of Concept Relationships F (87) = 0.314 *p*=ns	Section 001, Section 101, 002	Section 103, Section 104	0.06
* p < 0.05, ** p < 0.01, *** p < 0.001, (ns) = not significant			

When discussing the effect of feedback, it is important to mention that students might request and review feedback, and then paraphrase the suggested concept without using the exact same words. For instance, "Radical Change" was listed in the feedback as a suggested concept to include. However, “the entire backend support of IT cannot be radically changed…” was used in a submitted assignment. Another example was that “business transformation” was suggested as feedback, but the term was paraphrased as “the main techniques to transform the business…” in a submitted assignment. As a result, the number of concept relationships that could be identified by the system was likely affected as well, because the same concept could be expressed differently either by using synonyms of it or paraphrasing. That is why the numbers (experimental: 12 for Fall 2018 and 5 for Spring 2019; control: 1 for Fall 2018 and 2 for Spring 2019) seem to be very small. However, the overall picture was clear – the proposed system did have an effect on the number of concepts and concept relationships included in the submitted assignments. Therefore, providing students with immediate feedback while they are engaging with the writing tasks can encourage them to reflect and make improvement on their coursework to a greater extent.

In the proposed system, the ranked list of key concepts is displayed to students whenever they request feedback. Students are expected to learn to use the provided suggestions meaningfully in an assignment. If a student attempts to simply copy the suggested concepts and concept relationships into their assignment, it only affects the number of concepts and concept relationships suggested by the system in the following iteration when the student requests feedback again. In other words, the system can only provide feedback on what important concepts and relationships are missing, but not how they should be incorporated in an assignment. Students as active learners should decide how to apply these suggestions into their assignment on their own. On the other hand, students’ writing assignments are not evaluated based on how many concepts or concept relationships can be identified in their assignments. The number of concepts and concept relationships demonstrated in students' assignments can be used as one of the indicators of student mastery of conceptual knowledge, rather than the direct measurement of the assignment's quality and academic performance. Ultimately, the instructor will grade the assignment based on how well the assignment is written overall.

5.6 Discussion

The repeated measures study examined students' perceptions and utilization of the automated feedback, and the effects of the automated feedback on student learning in WTL activities. Based on a variation of UTAUT model, the survey instrument was designed based on constructs validated in prior research (Davis, 1989; Venkatesh et al., 2003) and adapted to the context of this research. Although previous studies suggest that the UTAUT model can be used to evaluate the users' acceptance of web-based learning environments, the focus of our research was to propose a technical framework to generate automated feedback to support student learning of conceptual knowledge in WTL activities and develop a proof-of-concept system to evaluate the generated automated feedback. Thus, the validation of the variation of UTAUT model used is beyond the scope of this study and thus not conducted. The major study results are discussed as follows.

Regarding students' perceptions of using the proposed system, the findings indicate that the participants hold positive attitudes toward the effects of using the automated feedback prototype system to support their learning in the process of writing an assignment. About 80% of students agreed or strongly agreed that the proposed automated feedback system is easy to use, user-friendly, stable, and easy to understand. Most participants agreed or strongly agreed that the automated feedback they received in the system was helpful for improving their coursework (78.4%) and motivating them to revise the assignment (77.3%). Furthermore, most students agreed or strongly agreed that the automated feedback they received from the proposed system was relevant (78.4%), easy to understand (82.9%), accurate (76.1%) and useful (79.5%). Survey results suggest that most of them (79.5%) would recommend doing this type of assignment with the proposed system for future classes.

As for students' utilization of the generated automated feedback, it was found that students tended to incorporate more concepts and concept relationships in their assignments when they received automated feedback as suggestions in the course of writing an assignment. Students' responses showed that most of them generally understood the purpose of the proposed system and how the automated feedback generated by the system could be used to support their writing and learning. As discussed, the main purpose of the proposed system is to support meaningful learning of conceptual knowledge by providing students with the most important and relevant domain concepts as suggestions for them to consider, rather than replacing instructors' feedback or scores for students' submitted assignment. Students' assignments will not be graded automatically by using the tool based just on the inclusion of the suggested concepts or concept relationships into their assignments. In a scenario where a student chooses to simply copy and paste the concepts from the feedback directly into his or her assignment, the instructor will provide the final feedback and grade it based on the actual writing quality eventually. Automated feedback can address some of the barriers faced by instructors regarding timely feedback, but the intended use of automated feedback systems is to complement and not replace instructors' feedback as suggested in previous study (Wilson & Andrada, 2016). Finally, students' creative inclusion of topics outside of lecture slides will not be penalized, as long as the assignment as a whole makes sense to the instructor. It is an instructor's responsibility to make the judgment on the quality of the writing assignments and provide assignment grading or additional feedback, as necessary.

5.7 Summary

This chapter discussed the data analysis methods and the results of the main study. Survey results show that most participants felt the feedback they received was relevant (78.4%), easy to understand (82.9%), accurate (76.1%) and useful (79.5%); they also felt that the proposed system made it easier to study course concepts (80.7%) and was useful in learning course concepts (77.3%). Based on the log analysis of students' actual usage of the system, all participants requested feedback at least once when they used the proposed system. After requesting feedback, 83 out of 88 participants revised their assignments. Analyses of students' submitted assignments reveal that more course concepts and concept relationships were included when they used the proposed system. Collectively, these results show that the proposed automated feedback prototype system contributes to students incorporating more course concepts and concept relationships into their writing assignments, as well as supporting their learning of conceptual knowledge in a WTL activity.

CHAPTER 6

DISCUSSION

This chapter summarizes the major findings of this study, discusses the limitations and contributions of this research. The chapter concludes by posting avenues for future work and implications for the development of an automated feedback system to support student learning in WTL activities.

6.1 Discussion of Research Findings

With the advancement of technology in education, there is a growing interest in developing better ways to facilitate and support student learning with automated feedback. The work in this dissertation proposed a new approach for automatic generation of feedback to assist student learning of conceptual knowledge in WTL activities, created an automated feedback prototype system, and conducted a series of studies to explore how students perceive and use the proposed system and how the automated feedback can influence student learning.

6.1.1 Key Research Findings

Based on the research questions this dissertation aims to investigate, the main findings from the evaluation are summarized as follows:

RQ1. *How effective is the automated formative feedback generated with the help of concept maps constructed from lecture slides and writing assignments?*

The effectiveness of feedback was evaluated using two methods in our studies. First, in our system, the domain knowledge representation constructed from instructors' lectures slides was used as a reference to generate automated feedback. Before the feedback can be generated, it is important to evaluate the system's ability to extract key concepts from instructors' lecture slides. Two independent domain experts were invited to participate in the evaluation. Our experimental results evaluated by Precision and Recall measures are comparable with those reported in related studies as discussed in Chapter 4.

Second, the automated feedback was tailored for a specific assignment draft written by individual students. We relied on the students' responses in the survey to inform us about the quality of the generated feedback. Based on the survey analysis, most participants (N=88) agreed or strongly agreed that the feedback they received in the proposed system was relevant (78.4%), easy to understand (82.9%), accurate (76.1%) and useful (79.5%), which are the main measures used to evaluate the effectiveness of feedback in related work (Roscoe, Allen, Johnson, & McNamara, 2018; Thurlings, Vermeulen, Bastiaens, & Stijnen, 2013).

RQ2. *Whether and how do students utilize the concept map-based formative feedback in WTL activities?*

In the proposed system, students can request feedback whenever they need some hints or suggestions. Based on the analysis of survey data, all the students self-reported that they requested the feedback in the course of writing the assignments. Only one student reported he or she did not read the feedback. When asked how carefully they read the feedback provided in the system, the majority of students across different course sections reported that they read the feedback carefully when they did the assignment.

More importantly, analysis of log data shows that all participants requested feedback at least once when they used the proposed system. After requesting feedback, 83 out of 88 participants revised their assignments. Based on the log analysis of students' actual usage of the system, during the fall semester of 2018, 29 out of 30 participants in Group 1 revised the assignment after requesting feedback for the first assignment, and 27 out of 30 in Group 2 made revisions after requesting feedback for the second assignment. During the spring semester of 2019, 13 out of 14 participants in Group 1 revised the assignment after requesting feedback for the first assignment, and 14 out of 14 in Group 2 made revisions after requesting feedback for the second assignment. In the following section 6.1.2, a case study illustrates how the prototype system can be used to its fullest potential in support student writing and learning during a WTL activity.

RQ3. *To what extent can the concept map-based formative feedback on writing assignments affect student learning outcomes?*

Learning outcomes in our study are measured in two dimensions: perceived learning and actual learning. Perceived learning is defined as the degree to which a person believes that using a particular system improves his or her learning, which was captured through post-assignment surveys. Survey results suggest that there was no statistically significant difference between the two systems. However, when asked to what extent the automated feedback helped students learn, most participants agreed or strongly agreed that the provided automated feedback can motivate them to revise their assignment (77.3%), gain a better understanding of course concepts (83%), and improve their assignment (78.4%).

Actual Learning refers to students' assignment performance, which can be captured by instructors' assignment grades, for students who consented to release their assignment grades. After analyzing instructors' assignments grades, we found them to be homogeneous in our study, and thus were not indictive of how the generated feedback affected learning performance. Previous studies suggest that the only way to determine if learning results from feedback is for students to utilize the feedback in the assignment writing process to complete the feedback loop, i.e., to ensure that information is received and acted on (Boud & Molloy, 2013; Sadler, 1989). Thus, actual learning was examined through log analysis of students' actual system usage and assignment performance. Analyses of students' submitted assignments completed during the main study across two semesters reveal that more course concepts and concept relationships were included when they used the proposed system. The differences between the mean values of the number of concepts and concept relationships shown in the assignments completed using the Write-and-Learn system and those of the number of concepts and concept relationships shown in the assignments completed using the baseline system were statistically significant. The proposed system did have an effect on the number of concepts and concept relationships included in the submitted assignments.

RQ4. *Do students recommend future use of the proposed automated feedback system? Why or why not?*

Although there were no statistically significant differences on the preferences for the two systems, the proposed system had a higher mean than the baseline system as the preferred system. Recommendation for use showed an interaction effect with the students' assigned system at the 90% confidence level. The majority (79.5%) of students would recommend doing this type of assignment with the proposed system for future classes. Survey results suggest that most students give positive comments concerning the usability of the proposed system. As for the perceived ease of use, more than 80% of students agreed or strongly agreed that the proposed system is easy to use, user-friendly, stable, and easy to understand. In terms of the perceived usefulness, two of the four items resulted in statistically significant differences between the proposed system and the baseline system: using the system would make it easier to study course concepts and the system is useful in learning course concepts. Also, perceived usefulness showed an interaction effect with the students' assigned system at the 90% confidence level. This result could be reflection of the capability of the proposed system to support student learning of course concepts in a WTL activity. Participants provided their opinions on what they liked best or least about the proposed system through open-ended questions as discussed in Chapter 5.

6.1.2 Case Study

The writing cycle typically consists of the following steps: write assignment, request feedback, review feedback, revise assignment and submit assignment. Each student follows his or her writing process. The advantage of a computer-based automated feedback system like Write-and-Learn is that the feedback can be delivered instantly and can be summoned as often as necessary, and can engage students in a cycle of writing, reviewing feedback, to revising their assignment.

In this section, a case study is presented to show how the prototype system can be used to its fullest potential in helping students improve their assignment and supporting their learning of course concepts throughout a WTL activity. During the studies, the system inputs of each participant including the log data, assignment drafts and finished assignments can be automatically collected. When we went through the data to understand user behaviors, several interesting cases were identified. This specific case study was identified from the pilot because a student made multiple attempts to request feedback and revised his or her assignment multiples times after the feedback was requested.

After the student logs into the system, he or she can read the assignment instructions, and then start with the first line before feedback can be requested. In the case study as shown in Table 6.1, the student requested feedback after completing the first line of a sentence. Once the student requested feedback, the system suggested a list of key concepts or relevant concepts based on what has been written for the student to consider including in the assignment. For the convenience of our discussion, five sets of suggestions are listed in the tables of case study below and they are separated by numbers, e.g., "1)" refers to set 1 of automated feedback. In the real system interface as illustrated in Chapter 3, the system only displays a set of top 10 suggestions in the first page, including missing concepts and unestablished concept relationships based on what the student writes. If the student would like to view more feedback, he or she can click the next pages for more suggestions. As there was no matching concept in the first draft as displayed in Table 6.1, only a list of key concepts was suggested to the student. Subsequently, the student can request feedback at any time during the process of writing an assignment and then revise the assignment as many times as necessary before the student submits it for grading.

Table 6.1 Case Study: First Draft

Student's Written Text	Automated Formative Feedback
Chapter 5 starts by comparing the SILO perspective with the business perspective.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; radical change; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; Sloan Valve; SAP; agile processes 2) Sub-optimization reinvent; Six Sigma; GE; Silo Functional Perspective; Adoption Decision; Business Transformation; customer requirements; Enterprise Resource Planning; Radical process; natural extension 3) favorable reactions; continuous process; aggressive improvement; <u>overall organizational objectives</u>; diverse marketplace; Process Perspective; cross-functional organization; Radical Change; Current systems; competitive advantage 4) Radical Change Approaches; external stakeholders; integrated supply; Mapping Processes; Organizational processes; Total Quality Management; Integrated Supply Chains; strategic advantage; <u>business processes</u>; Specialized sales 5) Development Process; Supply Chain Management; key metrics; Customer Relationship Management; breakthrough impact; Incremental Change; <u>organizational goals</u>; Common systems; small buffer; Resource Planning

After reviewing the feedback, the student wrote more content as shown in blue in Table 6.2, where three more concepts as underlined were included based on the suggestions provided by the system. For these three matching concepts that were identified from the student's draft, the system can identify other key concepts that are associated with them and then suggest to the student to consider establishing the relationships between them.

Table 6.2 Case Study: Second Draft

Student's Written Text	**Automated Formative Feedback**
Chapter 5 starts by comparing the SILO perspective with the business perspective. The Silo perspective has specialized functions and allows optimization of expertise, but it tends to lose sight of overall organizational objectives, has cross functional inefficiencies and leads to redundancy of information. On the other hand, in business processes perspective the focus is cross functional and on optimization of organizational goals and avoids work duplication but requires sophisticated software and knowledgeable generalists.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; radical change; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; Sloan Valve; SAP; agile processes 2) Sub-optimization reinvent; Six Sigma; GE; Silo Functional Perspective; Adoption Decision; Business Transformation; customer requirements; Enterprise Resource Planning; Radical process; natural extension 3) favorable reactions; continuous process; aggressive improvement; diverse marketplace; Process Perspective; cross-functional organization; Radical Change; Current systems; competitive advantage; Radical Change Approaches 4) external stakeholders; integrated supply; Mapping Processes; Organizational processes; Total Quality Management; Integrated Supply Chains; strategic advantage; Specialized sales; Development Process; Supply Chain Management 5) key metrics; Customer Relationship Management; breakthrough impact; Incremental Change; Common systems; small buffer; Resource Planning *Suggested Concept Relationships You Might Consider Establishing:* overall organizational objectives – Silo Functional Perspective business processes – Adoption decision

As shown in Table 6.3, the student chose to write more lines show in blue for his or her assignment draft. As no new matching concepts were included in the third draft this time, there was no change in the suggestions provided by the system.

Table 6.3 Case Study: Third Draft

Student's Written Text	Automated Formative Feedback
Chapter 5 starts by comparing the SILO perspective with the business perspective. The Silo perspective has specialized functions and allows optimization of expertise, but it tends to lose sight of overall organizational objectives, has cross functional inefficiencies and leads to redundancy of information. On the other hand, in business processes perspective the focus is cross functional and on optimization of organizational goals and avoids work duplication but requires sophisticated software and knowledgeable generalists. The chapter goes on to say that most businesses operate business processes, even if their organization charts are structured by functions rather than by processes. Organizations build dynamic business processes (with automatic updating) or agile business processes (easily reconfigurable) to stay competitive and consistently meet changing customer demands.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; radical change; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; Sloan Valve; SAP; agile processes 2) Sub-optimization reinvent; Six Sigma; GE; Silo Functional Perspective; Adoption Decision; Business Transformation; customer requirements; Enterprise Resource Planning; Radical process; natural extension 3) favorable reactions; continuous process; aggressive improvement; diverse marketplace; Process Perspective; cross-functional organization; Radical Change; Current systems; competitive advantage; Radical Change Approaches 4) external stakeholders; integrated supply; Mapping Processes; Organizational processes; Total Quality Management; Integrated Supply Chains; strategic advantage; Specialized sales; Development Process; Supply Chain Management 5) key metrics; Customer Relationship Management; breakthrough impact; incremental change; Common systems; small buffer; Resource Planning *Suggested Concept Relationships You Might Consider Establishing:* overall organizational objectives - Silo Functional Perspective business processes – Adoption decision

Then, in Table 6.4, the student started another paragraph by adding more concepts into the assignment draft. Several key concepts can be identified in the second paragraph. As more new matching concepts that can be identified in the draft, more suggestions that show the unestablished concept relationships are provided. For example, as the two concepts, "continuous process" and "Six Sigma", are closely related, one of them "continuous process" has been included in the written text, and the system recommends its relevant key concept "Six Sigma" for the student to consider establishing the relationship between them.

Table 6.4 Case Study: Fourth Draft

Student's Written Text	Automated Formative Feedback
Chapter 5 starts by comparing the SILO perspective with the business perspective. The Silo perspective has specialized functions and allows optimization of expertise, but it tends to lose sight of overall organizational objectives, has cross functional inefficiencies and leads to redundancy of information. On the other hand, in Business Processes perspective the focus is cross functional and on optimization of organizational goals and avoids work duplication but requires sophisticated software and knowledgeable generalists. The chapter goes on to say that most businesses operate business processes, even if their organization charts are structured by functions rather than by processes. Organizations build dynamic business processes (with automatic updating) or agile business processes (easily reconfigurable) to stay competitive and consistently meet changing customer demands. Business processes are transformed using two approaches, the incremental change and the radical change. The incremental change involves approaches such as Total Quality Management (TQM) which incorporates methods of continuous process improvement.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; Sloan Valve; SAP; agile processes; Sub-optimization reinvent 2) Six Sigma; GE; Silo Functional Perspective; Adoption Decision; Business Transformation; customer requirements; Enterprise Resource Planning; Radical process; natural extension; favorable reactions 3) aggressive improvement; diverse marketplace; Process Perspective; cross-functional organization; Radical Change; Current systems; competitive advantage; Radical Change Approaches; external stakeholders; integrated supply 4) Mapping Processes; Organizational processes; Integrated Supply Chains; strategic advantage; Specialized sales; Development Process; Supply Chain Management; key metrics; Customer Relationship Management; breakthrough impact 5) Common systems; small buffer; Resource Planning *Suggested Concept Relationships You Might Consider Establishing:* overall organizational objectives - Silo Functional Perspective business processes – Adoption Decision radical change – aggressive improvement continuous process – Six Sigma

As shown in Table 6.5, the student did take the suggestion by continuing the discussion of Six Sigma. As more concepts are incorporated into the assignment draft gradually, more suggestions are provided for the student to consider establishing the relationships between the relevant key concepts.

Table 6.5 Case Study: Fifth Draft

Student's Written Text	Automated Formative Feedback
Chapter 5 starts by comparing the SILO perspective with the business perspective. The Silo perspective has specialized functions and allows optimization of expertise, but it tends to lose sight of overall organizational objectives, has cross functional inefficiencies and leads to redundancy of information. On the other hand, in business processes perspective the focus is cross functional and on optimization of organizational goals and avoids work duplication but requires sophisticated software and knowledgeable generalists. The chapter goes on to say that most businesses operate business processes, even if their organization charts are structured by functions rather than by processes. Organizations build dynamic business processes (with automatic updating) or agile business processes (easily reconfigurable) to stay competitive and consistently meet changing customer demands. Business processes are transformed using two approaches, the incremental change and the radical change. The incremental change involves approaches such as Total Quality Management (TQM) which incorporates methods of continuous process improvement. Six Sigma is a data driven approach and methodology for eliminating defects from a process and has two processes namely DMAIC and DMADV. The radical change aims to make a rapid, breakthrough impact on key metrics and is used when major change is needed in a short time and as an example, the book illustrates how Sloan Valve improved the time-to-market of new products and reduced the time from 18–24 months down to 12 months.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; SAP; agile processes; Sub-optimization reinvent; GE 2) Silo Functional Perspective; Adoption Decision; Business Transformation; customer requirements; Enterprise Resource Planning; Radical process; natural extension; favorable reactions; aggressive improvement; diverse marketplace 3) Process Perspective; cross-functional organization; Radical Change; Current systems; competitive advantage; Radical Change Approaches; external stakeholders; integrated supply; Mapping Processes; Organizational processes 4) Integrated Supply Chains; strategic advantage; Specialized sales; Development Process; Supply Chain Management; Customer Relationship Management; Common systems; small buffer; Resource Planning *Suggested Concept Relationships You Might Consider Establishing:* overall organizational objectives - Silo Functional Perspective business processes – Adoption Decision radical change – aggressive improvement breakthrough impact – organizational goals Sloan Valve – Development Process

In the last attempt, the student wrote more by adding more concepts into the assignment draft. Overall, the student requested the feedback six times and then revised the assignment after reviewing the feedback each time. Analysis of the various drafts of this assignment shows that a total of 18 key course concepts were incorporated in the final assignment submission based on the automated feedback provided by the system, as displayed in Table 6.6. For some matching concepts, the student has recognized these concepts are associated with each other and discussed them together in the same paragraph. The system will not suggest to the student to consider establishing the relationships between them if these matching concepts can be identified in the same paragraph when the feedback is requested.

Table 6.6 Case Study: Final Submission

Student's Written Text	Automated Formative Feedback
Chapter 5 starts by comparing the SILO perspective with the business perspective. The Silo perspective has specialized functions and allows optimization of expertise, but it tends to lose sight of overall organizational objectives, has cross functional inefficiencies and leads to redundancy of information. On the other hand, in business processes perspective the focus is cross functional and on optimization of organizational goals and avoids work duplication but requires sophisticated software and knowledgeable generalists. The chapter goes on to say that most businesses operate business processes, even if their organization charts are structured by functions rather than by processes. Organizations build dynamic business processes (with automatic updating) or agile business processes (easily reconfigurable) to stay competitive and consistently meet changing customer demands. Business processes are transformed using two approaches, the incremental change and the radical change. The incremental change involves approaches such as Total Quality Management (TQM) which incorporates methods of continuous process improvement. Six Sigma is a data driven approach and methodology for eliminating defects from a process and has two processes namely DMAIC and DMADV. The radical change aims to make a rapid, breakthrough impact on key metrics and is used when major change is needed in a short time and as an example, the book illustrates how Sloan Valve improved the time-to-market of new products and reduced the time from 18–24 months down to 12 months. A workflow diagram is used to show a picture, or map, of the sequence and detail of each process step. Business Process Management systems are used to help managers design, control, and document business processes and ultimately workflow in an organization. Enterprise systems provide functionality needed to run a business and help share data between organizational divisions. They include four systems, the first one being Enterprise Resource Planning (ERP) systems which Seamlessly integrate information flows and are used to manage resources including financial, human resources and operations and allow external stakeholders to have access to information. The second one is Customer Relationship Management (CRM) and it is used to manage the processes related to customers and the relationships developed with customers for e.g., SAP, Oracle. The third one is the Supply Chain Management (SCM) system which manages integrated supply chain and streamlines planning and helps to deliver products to all members of the supply chain. And the fourth one is the Product Lifecycle Management (PLM). In the end when it comes to adoption decision, enterprise systems should drive business process design when starting out or when current systems are in crisis and not when the package doesn't fit or there is no top management support.	*Suggested Concepts You Might Consider Including:* 1) threatening change; Relationship Management; cross-functional process; detailed flow diagram; Key Aspects; near-perfect products; agile processes; Sub-optimization reinvent; GE; Silo Functional Perspective 2) Business Transformation; customer requirements; Radical process; natural extension; favorable reactions; aggressive improvement; diverse marketplace; Process Perspective; cross-functional organization; Radical Change 3) Current systems; competitive advantage; Radical Change Approaches; integrated supply; Mapping Processes; Organizational processes; strategic advantage; Specialized sales 4) Development Process; Common systems; small buffer; Resource Planning *Suggested Concept Relationships You Might Consider Establishing:* overall organizational objectives - Silo Functional Perspective business processes – Adoption Decision radical change – aggressive improvement breakthrough impact – organizational goals Sloan Valve – Development Process

Figure 6.1 illustrates the screenshot of the final written assignment submitted by the student after using the system. The case study demonstrates how the prototype system can be utilized to its fullest potential with the goal of supporting student learning of conceptual knowledge during the process of writing an assignment.

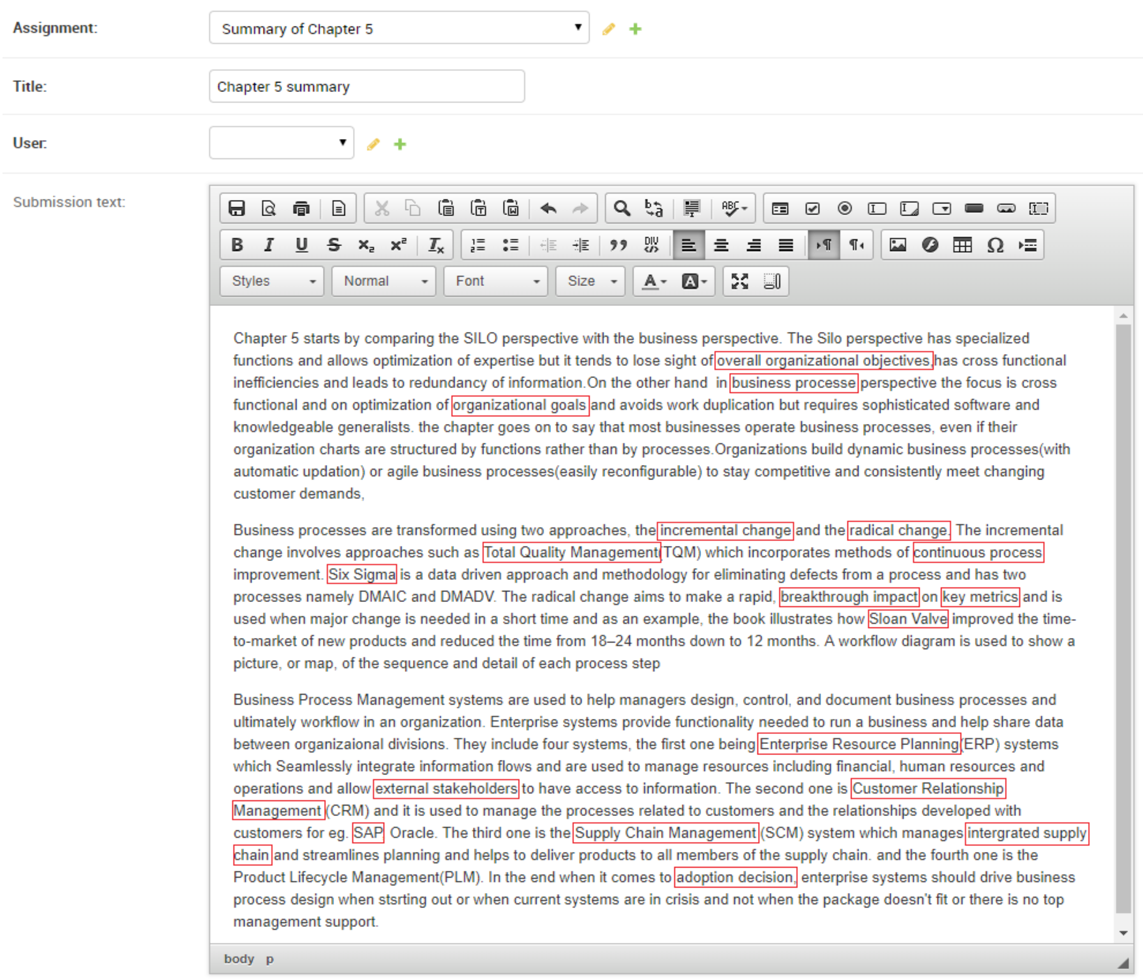

Figure 6.1 Case study.

As pointed out by (Forsman, 1985), learning allows a newly discovered rock to be fitted into the wall. It is impossible for instructors to forcibly pile all the new rocks into the students' minds. Each student uses his or her own mind to pick through the rocks, quickly or slowly, to rebuild or enlarge their knowledge structure. As present in this case study, if the student follows the suggestions, more concepts and concept relationships can be included in the writing assignment.

The underlying theory echoes the purpose of developing the automated feedback system, which is focused on providing students with the most important and relevant domain concepts as suggestions for them to consider, rather than being used as a grading tool to give students a numeric score for their submitted assignment. The quality of the final written assignment will be judged by the instructor; therefore, even if students did not incorporate the suggested concepts in a meaningful way, they would not receive credit just by randomly adding system suggested concepts. Using the system will not affect students' creativity if students would like to discuss additional concepts or ideas that are not covered in the lecture slides. Moreover, students are encouraged to assimilate the most important and relevant domain concepts into their existing knowledge structure to construct a more integrated knowledge structure. The automated feedback in the system is expected to help the student reflect more on the course concepts, motivate the student to assimilate new concepts into his or her existing knowledge structure, and hence revise the assignment or improve the coursework.

6.2 Limitations

This section presents the primary limitations of this research. As a proof of concept study, the major limitation of this work is generalizability, because only one graduate course was used in the evaluation. Another limitation of this research is related to the constraints of the writing assignments used in our main study. In the main study, students were asked to complete two writing assignments that were required by the course, which would encourage more serious participation and motivate them to interact with the system more, compared to the extra credit assignments used in the pilot study. However, these two writing assignments used in the main study were in the form of online discussions. Because the assigned systems could not support group activities, students still needed to post their summary as the initial comments in the discussion forum on Moodle for further discussion of the course topic, after they completed the summary in the assigned systems. It might have increased the complexity of the research task and caused confusion; especially when participants responded to the second survey days after completing the experiment to compare the baseline system and the proposed system, they might have been mistaken about Moodle being one of the systems.

6.3 Contributions

The findings of this dissertation contribute to the field of educational research in the following ways. First, to the best of our knowledge, this research is the first attempt to explore an automatic approach to generate actionable and individualized formative feedback based on knowledge representations extracted from instructors' lecture slides and individual students' writing assignments. The automatic approach for generating formative feedback proposed in this research demonstrates that it is feasible to generate feedback automatically based on the comparison of the representations of domain knowledge and student knowledge. The combination of theories and technologies creates new opportunities to address a wide gap in understanding of how to assist student learning of conceptual knowledge through automated feedback on writing assignments with NLP techniques.

Second, this work brings together what is known about cognitive sciences, learning sciences and assessment principles to address the research gap concerning how to assist student learning of conceptual knowledge in WTL activities by proposing an automated feedback system. The Write-and-Learn prototype system was implemented, which shows that the proposed methodology to derive automated feedback during the writing process is practically feasible. The benefits of the proposed automated feedback system include but are not limited to: (1) to provide support for a WTL activity that fosters learning of conceptual knowledge in the course of writing an assignment; (2) to give students extended practice in writing and revising assignments while relieving instructors from the burden of providing timely feedback and reviewing successive drafts; and (3) to motivate students to review the course content and reflect on their learning by providing real-time and individualized feedback on their assignment. The proposed system is well suited to a learning environment where domain concepts and their relationships can be extracted from textual course materials (e.g., lecture slides), student concepts and their relationships can be extracted from textual assignments, and formative feedback can be automatically generated during the writing tasks.

Finally, a series of research studies were conducted to provide empirical evidence on how and to what extent automated feedback can be used to assist student learning in WTL activities. These outcomes of the research activities can lay the groundwork for future studies in this area and be of significant value to the research community. Based on the survey results, students reported that they felt they benefited from the provided feedback and enjoyed using the proposed system. Such a feedback generation approach can be applied to other WTL assignments such as reflective essays and summary, etc. The system is not intended to replace the roles of instructors to evaluate the final submissions of students' writing assignments. This tool is intended to be used by students independently, and thereby to provide students with more opportunities in reflecting on their learning and improving their writing, while leaving instructors more time for other kinds of educational activities, such as preparing course materials, delivering instruction and evaluating final versions of students' assignments.

6.4 Future Work

Students and instructors play different roles when utilizing educational technology. The current prototype system can stimulate students' learning by making them aware of their own cognitive processes in WTL activities. As for instructors, they are engaged in efforts to improve instruction, measure the efficacy of different teaching practices, and understand to what extent students learn the concepts and practices that are fundamental to their disciplines. If instructors can be provided with timely feedback on progress towards students' mastery of content, they can gain a more detailed understanding of what their students know and can do, personalize learning pathways for their students, and strategically adjust instruction more quickly to meet the needs of diverse students (U.S. Department of Education, 2015).

One future direction of our study is to generate feedback for the instructors regarding how well students are doing in terms of learning the core domain knowledge -- by comparing all student knowledge representations generated from the completed assignments with the domain knowledge representation, the system can find out common missing core concepts and concepts relationships. The instructor can choose to immediately reinforce those concepts in the next lecture, while student learning is still taking placing rather than at the end of a course or a program. Such a tool would afford the instructors the opportunities to make immediate changes to classroom activities and assignments or give students prompt guidance.

6.5 Summary

Existing feedback systems for writing provide opportunities for students to write, receive feedback, and then revise writing products in a timely iterative cycle (Foltz & Rosenstein, 2015). However, such feedback is mainly focused on improving students' writing skills rather than their conceptual development. In the context of WTL, if individual students can be provided with formative feedback to help them acquire and develop their conceptual knowledge, rather than simply focusing on their writing skills, that would help them develop conceptual understanding, improve academic performance and achieve higher learning outcomes.

To make this a reality, this research proposes an automatic approach to generate formative feedback by utilizing the concept maps constructed from lecture slides and writing assignments. Building on existing work in the area of concept mapping, the automated feedback in our approach can suggest a list of key concepts and unestablished concept relationships for students to consider during the process of writing an assignment. Besides, the automated feedback prototype system was developed with a goal of supporting meaningful learning of conceptual knowledge in WTL activities. Furthermore, the empirical studies were conducted to investigate the effects of the automated feedback on student learning, understand student perceptions of the automated feedback prototype system, and examine the use of the generated formative feedback. With this work, we hope to promote meaningful learning of conceptual knowledge through automated formative feedback during a WTL activity and gain insights into how such feedback can be better provided to satisfy the needs of students.

APPENDIX A

CONSENT FORMS AND FERPA WAIVER

Students in participating courses were invited to participate in the research study in return for extra credit. Students who do not wish to participate in the research will be offered an equivalent alternative assignment of the instructor's choice to obtain the same amount of extra credit.

To indicate their agreement to participate in the research, students are asked to complete and sign the consent form shown in this appendix. An optional Family Educational Rights and Privacy Act (FERPA) waiver is included in the consent form. Students who sign the FERPA waiver consent to release their assignment grades to the researcher for research purposes.

New Jersey Institute of Technology
323 Martin Luther King Blvd.
Newark, NJ 07102

CONSENT TO PARTICIPATE IN A RESEARCH STUDY

TITLE OF STUDY:
Write-and-Learn Study

RESEARCH STUDY:
I, ________________________________, have been asked to participate in a research study under the direction of Dr. Yi-fang Brook Wu and Ye Xiong. Other professional persons who work with them as study staff may assist.

PURPOSE:
The purpose of this Write-and-Learn study is to explore how to provide students with automated feedback to support their learning of conceptual knowledge during the process of writing an assignment.

DURATION:
My participation in this study will last for one and half months.

PROCEDURES:
I have been told that, during the course of this study, the following will occur:

I will sign this consent form and FERPA form indicating my willingness to participate in the study. I will complete the writing assignment with the assigned system. The assignment will be written and submitted using the assigned system. When I am logged in to the system, the information I enter into the system will be logged for analysis. I will be asked to complete a brief online survey prior to beginning the assignment and another brief online survey upon completion of the assignment. The online surveys will be administered through SurveyMonkey (an online survey tool).

PARTICIPANTS:
I will be one of about 90 participants in this study.

RISKS/DISCOMFORTS:
Online surveys are conducted on a secure (https) server. As an online participant, there is always the risk of intrusion by outside agents (i.e., hacking) and, therefore the possibility of being identified exists. Every reasonable effort will be made to minimize the risk. All data will be stored in a secure NJIT computer server and will not be made available to anyone except Dr. Yi-fang Brook Wu (PI) and Ye Xiong (Co-PI) who are involved in this research.

There also may be risks and discomforts that are not yet known. I fully recognize that there are risks that I may be exposed to by volunteering in this study which are inherent in participating in any study; I understand that I am not covered by NJIT's insurance policy for any injury or loss I might sustain in the course of participating in the study.

CONFIDENTIALITY:
I understand confidential is not the same as anonymous. Confidential means that my name will not be disclosed if there exists a documented linkage between my identity and my responses as recorded

in the research records. Every effort will be made to maintain the confidentiality of my study records. If the findings from the study are published, I will not be identified by name. My identity will remain confidential unless disclosure is required by law.

PAYMENT FOR PARTICIPATION:
I have been told that I will receive extra credit for my participation in this study.

RIGHT TO REFUSE OR WITHDRAW:
I understand that my participation is voluntary and I may refuse to participate, or may discontinue my participation at any time with no adverse consequence. I also understand that the investigator has the right to withdraw me from the study at any time.

INDIVIDUAL TO CONTACT:
If I have any questions about my treatment or research procedures, I understand that I should contact the principal investigator at:

Dr. Yi-fang Brook Wu
Informatics Department
New Jersey Institute of Technology
(973) 596-3368
yi-fang.wu@njit.edu

Ye Xiong
Informatics Department
New Jersey Institute of Technology
(973) 596-3368
yx98@njit.edu

If I have any addition questions about my rights as a research subject, I may contact:

Horatio Rotstein, IRB Chair
New Jersey Institute of Technology
323 Martin Luther King Boulevard
Newark, NJ 07102
(973) 596-5825
irb@njit.edu/ farzan@njit.edu

SIGNATURE OF PARTICIPANT
I have read this entire form, or it has been read to me, and I understand it completely. All of my questions regarding this form or this study have been answered to my complete satisfaction. I agree to participate in this research study.

Participant Name ______________________________

NJIT Email Address ______________________________

Signature ______________________________

Date ______________________________

AUTHORIZATION TO RELEASE INFORMATION

Student Consent for Educational Records to be Released to Researcher for Purposes of Analysis Only

Student's Name (please print): ______________________________ **NJIT UCID:** ______________________________

PLEASE READ:

In accordance with the Family Educational Rights and Privacy Act of 1974 (FERPA), the undersigned student hereby permits New Jersey Institute of Technology to disclose the information specified below to the researchers, Dr. Yi-fang Brook Wu and Ye Xiong, for the purposes of research only. This information will be kept strictly confidential and will not be disclosed to any third parties, nor will any identifiable information about the student be released.

This consent shall be valid for the semester during which the student participates in the research study.

INFORMATION TO BE RELEASED:

The information to be obtained shall be limited to:

- The student's grade for the course assignments designated as part of this research

I have read and understand the contents of this consent form pertaining to the Family Educational Rights and Privacy Act of 1974.

Student's Signature: ______________________________ **Date:** ______________________________

APPENDIX B

ASSIGNMENT INSTRUCTIONS

The assignment instructions for all research assignments are shown here.

B.1 Assignment Instructions for Pilot Study

Instructions

Write a summary of the chapter 5 “IT and Business Transformation”. Summarize the chapter in two - three paragraphs. Discuss the concepts that you consider important and why.

B.2 Assignment Instructions for Main Study

Instructions

1. Critique/analysis of the article - IT Doesn't Matter

Write your initial comment for “Discussion - IT Doesn’t Matter” based on the following article: *Carr, N. G. (2003). IT doesn't matter. Educause Review, 38, 24-38.*

The initial comment is to write a summary and address the questions posted as below (worth a maximum 6 points). It must be a minimum of 3 paragraphs (5-6 sentences each paragraph). Your initial comment must include the following:
1) What is the purpose of this paper?
2) Why is this important to us as IT professionals?
3) Does Carr's arguments seem reasonable, what are they and how would you argue against it? What would you argue to show he is wrong?
4) Critique/analysis of the article; relate this to the lecture.

2. Critique/analysis of the article - Cloud Computing

Write your initial comment for “Discussion - Cloud Computing” based on the following article: *Iyer, B. and Henderson, J.C. 2012. "Business value from clouds: Learning from users," MIS Quarterly Executive, 11/1, pp. 51-60.*

The initial comment is to write a summary and address the questions posted as below (worth a maximum 6 points). It must be a minimum of 3 paragraphs (5-6 sentences each paragraph). Your initial comment must include the following:
1) What is the purpose of this paper?
2) Will industry follow this on cloud computing? If so, why?

3) Critique/analysis of the article; relate this to the lecture.

3. Critique/analysis of the article - Big Data

Write your initial comment for “Discussion - Big Data” based on the following article: *McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. Harvard Business Review, 90(10), 60-68.*

The initial comment is to write a summary and address the questions posted as below (worth a maximum 6 points). It must be a minimum of 3 paragraphs (5-6 sentences each paragraph). Your initial comment must include the following:

1) What is the purpose of this paper?
2) Will industry follow this on big data? If so, why?
3) Critique/analysis of the article; relate this to the lecture.

APPENDIX C

RESEARCH SURVEYS

The survey questionnaires for all the studies are shown here.

C.1 Pre-Assignment Survey for Pilot Study

Before students started to write the assignment using the assigned system, they were asked to complete the pre-assignment survey. The questions included in the pre-assignment survey for the pilot study are shown as below.

Table C.1 Pre-Assignment Survey for Pilot Study

1	Participant Name (First Name and Last Name)
2	UCID
3	Gender (M/F)
4	Age
5	Level of Study (Undergraduate/Graduate Student)
6	Academic Program
	Computer Science Information Systems Information Technology Electrical and Computer Engineering Engineering Technology Management Other
7	Student Type (Domestic/International)
8	Is English your first language? (Yes/No)
9	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	Feedback is important to me.
	Feedback can guide me to improve my coursework.
	Feedback motivates me to study.
	Feedback tells me what I need to do to improve my performance in a subject.
	It is more important for me to know how to improve my coursework before I receive a particular grade.
	I always read the feedback on my assignments.
	I always use the feedback to improve my assignments.
10	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)

	Individualized feedback is better because it can be applied to improve my coursework.
	Real-time feedback is better because it helps me to immediately find out if I need to make appropriate changes.
	Specific feedback is better because it helps me to understand what I did right and wrong in an assignment.
	Relevant feedback is better because it gives me clear instructions for how to revise my assignment.
11	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I receive enough feedback from my instructors.
	My instructors provide enough information to make feedback useful.
	My instructors are always willing to provide feedback.
	The feedback I receive is relevant to my goals as a student.
	The feedback I receive is related to the purpose of the assignment.
12	Please indicate how often you receive the following types of feedback: (Never/Less Than Once a Month/Once a Month/Several Times a Month/Once a Week/Several Times a Week/Several Times a Day)
	Individual verbal feedback from the instructor Individual written comments from the instructor Group verbal feedback from the instructor Group written comments from the instructor Peer feedback Self-assessment Automated feedback
13	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I feel confident writing assignments in English.
	I have the necessary skills for writing assignments in English.
14	How easy/difficult do you find this course? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)
15	How easy/difficult do you find this writing assignment? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)

C.2 Post-Assignment Survey 1 for Pilot Study

After students had completed the required assignment using the system without feedback, they were asked to complete the post-assignment survey 1. The questions included in the post-assignment survey 1 for the pilot study are shown as below.

Table C.2 Post-Assignment Survey 1 for Pilot Study

1	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The grading rubric is fair and explicit enough.
	The assignment instruction is clear enough.
	When I need help to use the system, guidance is available to me.
	Special instruction concerning the system use is available to me.
	A specific person/group is available for assistance with any difficulties related with the system use.
2	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The system is easy to use.
	The system is user-friendly.
	The operation of the system is stable.
	My interaction with the system is clear and understandable.
	Interacting with the system does not require a lot of my mental effort.
	I find it easy to get the system to do what I want it to do.
3	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	Using the system would improve my learning performance.
	Using the system would increase my academic productivity.
	Using the system would make it easier to study course concepts and academic content.
	I find the system to be useful in learning course concepts and academic content.
4	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I developed a good understanding of the basic concepts during this assignment.
	The writing assignment helped me learn some important concepts in this course.
	I learned to see relationships between course concepts during this assignment.
	I developed an improved ability to integrate course concepts through the writing assignment.
	The writing assignment helped me gain a better understanding of the subject matter and course content.
5	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)

	I would recommend the system to my friends and classmates.
6	What is your overall impression of the system? Name three words or characteristics that describe the system.
7	What do you like about the system? What are the three things you like best about the system?
8	What do you dislike about the system? What are the three things you like least about the system?

C.3 Post-Assignment Survey 2 for Pilot Study

After students had completed the assignment using the system with feedback, they were asked to complete the post-assignment survey 2. The questions included in the post-assignment survey 2 for the pilot study are shown as below.

Table C.3 Post-Assignment Survey 2 for Pilot Study

1	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The grading rubric is fair and explicit enough.
	The assignment instruction is clear enough.
	When I need help to use the system, guidance is available to me.
	Special instruction concerning the system use is available to me.
	A specific person/group is available for assistance with any difficulties related with the system use.
2	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The system is easy to use.
	The system is user-friendly.
	The operation of the system is stable.
	My interaction with the system is clear and understandable.
	Interacting with the system does not require a lot of my mental effort.
	I find it easy to get the system to do what I want it to do.
3	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	Using the system would improve my learning performance.
	Using the system would increase my academic productivity.
	Using the system would make it easier to study course concepts and academic content.
	I find the system to be useful in learning course concepts and academic content.
4	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I developed a good understanding of the basic concepts during this assignment.
	The writing assignment helped me learn some important concepts in this course.
	I learned to see relationships between course concepts during this assignment.
	I developed an improved ability to integrate course concepts through the writing assignment.
	The writing assignment helped me gain a better understanding of the subject matter and course content.
5	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)

	The content of feedback is relevant.
	The content of feedback is easy to understand.
	The content of feedback is accurate.
	The content of feedback is useful.
6	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I enjoy receiving the feedback while I am writing the assignment.
	I am motivated to do my best work.
	I enjoy the writing and learning process in the system.
	I find using the system to be enjoyable.
	The actual process of using the system is pleasant.
	I have fun using the system.
7	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The feedback motivated me to learn more.
	The feedback encouraged me to reflect on how I can improve my assignments.
	The feedback helped me find out what I do not know so I could learn it.
	The feedback helped me improve the quality of my assignment.
8	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I would recommend the system to my friends and classmates.
9	What is your overall impression of the system? Name three words or characteristics that describe the system.
10	What do you like about the system? What are the three things you like best about the system?
11	What do you dislike about the system? What are the three things you like least about the system?

C.4 Pre-Assignment Survey for Main Study

Before students started to write the first required assignment using the assigned system, they were asked to complete the pre-assignment survey. The questions included in the pre-assignment survey for the main study are shown as below.

Table C.4 Pre-Assignment Survey for Main Study

1	Name (First Name and Last Name)
2	UCID
3	Gender Identity (M/F/Other)
4	Age
	18-20 21-24 25-29 30-39 40-49 >50
5	Level of Study (Undergraduate/Graduate)
6	Degree program
	Computer Science Information Systems Information Technology Electrical and Computer Engineering Engineering Technology Management Other
7	Student Type (Domestic/International)
8	Is English your first language? (Yes/No)
9	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	Feedback is important to me.
	Feedback can guide me to improve my coursework.
	Feedback motivates me to study.
	Feedback tells me what I need to do to improve my performance in a subject.
	It is more important for me to know how to improve my coursework before I receive a particular grade.
	I always read the feedback on my assignments.
	I always use the feedback to improve my assignments.
10	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I receive enough feedback from my instructors.
	My instructors are always willing to provide feedback.
	My instructors provide enough information to make feedback useful.
	The feedback I receive is relevant to my goals as a student.

	The feedback I receive is related to the purpose of the assignment.
11	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I feel confident writing assignments in English.
	I have the necessary skills for writing assignments in English.
	I feel confident finding information in the system.
	I have the necessary skills for using the system.
	I feel confident using system features.
	I feel confident operating system functions.
12	Please indicate how often you receive the following types of feedback: (Never/Less Than Once a Month/Once a Month/Several Times a Month/Once a Week/Several Times a Week/Several Times a Day)
	Individual verbal feedback from the instructor Individual written comments from the instructor Group verbal feedback from the instructor Group written comments from the instructor Peer feedback Self-assessment Automated feedback
13	How easy/difficult do you find this course? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)
14	How easy/difficult do you find this writing assignment? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)

C.5 Post-Assignment Survey 1 for Main Study

After students had completed the required assignment using the system without feedback, they were asked to complete the post-assignment survey 1. The questions included in the post-assignment survey 1 for the main study are shown as below.

Table C.5 Post-Assignment Survey 1 for Main Study

1	Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The grading rubric is fair and explicit enough.
	The assignment instruction is clear enough.
	When I need help to use the system, guidance is available to me.
	Special instruction concerning the system use (e.g., system demo, system instructions) is available to me.
	A specific person/group is available for assistance with any difficulties related with the system use.
2	The system is easy to use.
	The system is user-friendly.
	The operation of the system is stable.
	My interaction with the system is clear and understandable.
	Interacting with the system does not require a lot of my mental effort.
3	Using the system would improve my learning performance.
	Using the system would increase my academic productivity.
	Using the system would make it easier to study course concepts.
	I find the system to be useful in learning course concepts.
4	The writing assignment helped me learn some important concepts in this course.
	I learned to see relationships between course concepts through the writing assignment. (For example, “Six Sigma” is closely related to “Total Quality Management”.)
	I developed an improved ability to integrate course concepts through the writing assignment.
5	I would recommend doing the assignment with this system for future classes.
6	What is your overall impression of the system? Name three words or characteristics that describe the system.
7	What are the three things you like best about the system?
8	What are the three things you like least about the system?

C.6 Post-Assignment Survey 2 for Main Study

After students had completed the required assignment using the system with feedback, they were asked to complete the post-assignment survey 2. The questions included in the post-assignment survey 2 for the main study are shown as below.

Table C.6 Post-Assignment Survey 2 for Main Study

1	Related to your completion of the recent assignment with this version of the Write and Learn system, please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	The grading rubric is fair and explicit enough.
	The assignment instruction is clear enough.
	When I need help to use the system, guidance is available to me.
	Special instruction concerning the system use (e.g., system demo, system instructions) is available to me.
	A specific person/group is available for assistance with any difficulties related with the system use.
2	The system is easy to use.
	The system is user-friendly.
	The operation of the system is stable.
	My interaction with the system is clear and understandable.
	Interacting with the system does not require a lot of my mental effort.
3	Using the system would improve my learning performance.
	Using the system would increase my academic productivity.
	Using the system would make it easier to study course concepts.
	I find the system to be useful in learning course concepts.
4	The writing assignment helped me learn some important concepts in this course.
	I learned to see relationships between course concepts through the writing assignment. (For example, “Six Sigma” is closely related to “Total Quality Management”.)
	I developed an improved ability to integrate course concepts through the writing assignment.
5	I would recommend doing the assignment with this system for future classes.
6	The content of feedback is relevant.
	The content of feedback is easy to understand.
	The content of feedback is accurate.
	The content of feedback is useful.
7	I enjoy receiving the feedback while I am writing the assignment.
	I am motivated to do my best work.
	I enjoy the writing and learning process in the system.
	I find using the system to be enjoyable.

8	Did you request feedback from the system while writing your assignment? (Yes/No)
9	Did you read the feedback provided in the Write-and-Learn system? (Yes/No)
10	How carefully did you read the feedback provided in the Write-and-Learn system? (Extremely Carefully/Very Carefully/Somewhat Carefully/Not So Carefully/Not at All Carefully)
11	Did you revise your assignment based on the provided feedback? (Yes/No)
12	What is your overall impression of the system? Name three words or characteristics that describe the system.
13	What are the three things you like best about the system?
14	What are the three things you like least about the system?

C.7 Post-Assignment Survey 3 for Main Study

After students had completed two required assignments using the assigned systems, one with feedback and one without feedback, we would like to ask them to compare two systems and assess the extent to which they found the feedback to be of value. The questions included in the post-assignment survey 3 were attached to the second post-assignment survey for the main study.

Table C.7 Post-Assignment Survey 3 for Main Study

1	You have now completed two assignments using the assigned systems, one with automated feedback available and one without. We would like you to assess the extent to which you found the feedback to be of value. Please indicate your level of agreement with the following statements. (Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree)
	I found the feedback to be a waste of time.
	I found the feedback to be irrelevant to my writing assignment.
	I found the feedback to be useless to my learning.
	I found the feedback to be hard to understand.
	I found the feedback to be valuable to my learning.
	I preferred the No Feedback version of the system.
2	I learned from the feedback during the process of writing the assignment.
	The feedback motivated me to revise my assignment.
	The feedback encouraged me to reflect on how I could improve my assignment.
	The feedback helped me gain a better understanding of course concepts.
	The feedback helped me find out what I do not know.
	The feedback helped me improve the quality of my assignment.
3	This was the second assignment of this type which you completed for this class. Compare your experience in the second time to the first time you did this type of assignment and respond to the following statements: (A Great Deal Less/Slightly Less/About the Same/Slightly More/A Great Deal More)
	I liked the second system…
	Comments
4	What I liked about the system with feedback was…?
5	What I disliked about the system with feedback was…?

APPENDIX D

RESEARCH VARIABLES

Table D.1 Definitions of Research Variables

Variables	Definition
Self-Efficacy (SE)	The belief that one has the capability to perform a particular behavior, including system-specific self-efficacy and writing self-efficacy.
Effort Expectancy (EE)	A person's perceived course difficulty and assignment complexity.
Facilitating Conditions (FC)	Perceived enablers or barriers in the environment that influence a person's perception of ease or difficulty of performing a task. Including the clarity, perceived quality, and fairness of the assignment, grading rubrics and procedures.
Perceived Feedback Effectiveness (PFE)	A person's perception of the quality and effectiveness of the provided feedback.
Perceived Enjoyment (PE)	The degree to which a person believes that using a particular system would be enjoyable. An adaption of intrinsic motivation.
Perceived Ease of Use (PEU)	The degree to which a person believes that using a particular system would be free from effort.
Perceived Usefulness (PU)	The degree to which a person believes that using a particular system would enhance his or her academic performance.
Perceived Learning (PL)	The degree to which a person believes that using a particular system would improve his or her learning.
Actual Learning (AL)	Students' assignment performance.
Recommendation for Use (ROU)	Substitute for "Intention to Use". The degree to which a person would recommend the system to be used in the future.

Table D.2 Measures of Research Variables

Variables	Label	Items
Self-Efficacy (SE)	SE_1	I feel confident writing assignments in English.
	SE_2	I have the necessary skills for writing assignments in English.
	SE_3	I feel confident finding information in the system.
	SE_4	I have the necessary skills for using the system.
	SE_5	I feel confident using system features.
	SE_6	I feel confident operating system functions.
Effort Expectancy (EE)	EE_1	How easy or difficult do you find this course? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)
	EE_2	How easy or difficult do you find this writing assignment? (Very Easy/Easy/Neither Easy nor Difficult/Difficult/Very Difficult)
Facilitating Conditions (FC)	FC_1	The grading rubric is fair and explicit enough.
	FC_2	The assignment instruction is clear enough.
	FC_3	When I need help to use the system, guidance is available to me.
	FC_4	Special instruction concerning the system use (e.g., system demo, system instructions) is available to me.
	FC_5	A specific person/group is available for assistance with any difficulties related with the system use.
Perceived Feedback Effectiveness (PFE)	PFE_1	The content of feedback is relevant.
	PFE_2	The content of feedback is easy to understand.
	PFE_3	The content of feedback is accurate.
	PFE_4	The content of feedback is useful.
Perceived Enjoyment (PE)	PE_1	I enjoy receiving the feedback while I am writing the assignment.
	PE_2	I am motivated to do my best work.
	PE_3	I enjoy the writing and learning process in the system.
	PE_4	I find using the system to be enjoyable.
Perceived Ease of Use (PEU)	PEU_1	The system is easy to use.
	PEU_2	The system is user-friendly.
	PEU_3	The operation of the system is stable.
	PEU_4	My interaction with the system is clear and understandable.
	PEU_5	Interacting with the system does not require a lot of my mental effort.
Perceived Usefulness (PU)	PU_1	Using the system would improve my learning performance.
	PU_2	Using the system would increase my academic productivity.

	PU_3	Using the system would make it easier to study course concepts.
	PU_4	I find the system to be useful in learning course concepts.
Perceived Learning (PL)	PL_1	The writing assignment helped me learn some important concepts in this course.
	PL_2	I learned to see relationships between course concepts through the writing assignment. (For example, "Six Sigma" is closely related to "Total Quality Management".)
	PL_3	I developed an improved ability to integrate course concepts through the writing assignment.
	PL_4	I learned from feedback during the process of writing the assignment.
	PL_5	The feedback motivated me to revise my assignment.
	PL_6	The feedback encouraged me to reflect on how I can improve my assignments.
	PL_7	The feedback helped me gain a better understanding of course concepts.
	PL_8	The feedback helped me find out what I do not know.
	PL_9	The feedback helped me improve the quality of my assignment.
Actual Learning (AL)	AL	Students' assignment performance
Recommendation for Use (ROU)	ROU	I would recommend doing the assignment with this system for future classes.

REFERENCES

Aguiar, C. Z., Cury, D., & Zouaq, A. (2016). *Automatic Construction of Concept Maps from Texts*. Paper presented at the Seventh International Conference on Concept Mapping, Tallinn, Estonia.

Al-Assaf, N., Almarabeh, T., & Eddin, L. N. (2015). A Study on the Impact of Learning Management System on Students of the University of Jordan. *Journal of Software Engineering and Applications, 8*(11), 590.

Al-Shaer, I. M. R. (2014). Employing Concept Mapping as a Pre-writing Strategy to Help EFL Learners Better Generate Argumentative Compositions. *International Journal for the Scholarship of Teaching and Learning, 8*(2).

Anderson, J. R. (1993). Problem Solving and Learning. *American Psychologist, 48*(1), 35.

Andrews, K., Wohlfahrt, M., & Wurzinger, G. (2009). *Visual Graph Comparison.* Paper presented at the 13th International Conference Information Visualisation, Barcelona, Spain.

Atapattu, T., Falkner, K., & Falkner, N. (2012). *Automated Extraction of Semantic Concepts from Semi-structured Data: Supporting Computer-based Education through the Analysis of Lecture Notes.* Paper presented at the 23rd International Conference on Database and Expert Systems Application, Vienna, Austria.

Atapattu, T., Falkner, K., & Falkner, N. (2014a). *Acquisition of Triples of Knowledge from Lecture Notes: A Natural Language Processing Approach*. Paper presented at the 7th International Conference on Educational Data Mining, London, United Kingdom.

Atapattu, T., Falkner, K., & Falkner, N. (2014b). *Evaluation of Concept Importance in Concept Maps Mined from Lecture Notes: Computer vs Human*. Paper presented at the 6th International Conference on Computer Supported Education, Barcelona, Spain.

Atkinson, R. D., & Mayo, M. J. (2010). *Refueling the US Innovation Economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education*. Washington, DC: Information Technology and Innovation Foundation.

Attali, Y., & Burstein, J. (2004). Automated Essay Scoring with e-Rater® v. 2.0. *ETS Research Report Series, 2004*(2).

Attuquayefio, S., & Addo, H. (2014). Review of Studies with UTAUT as Conceptual Framework. *European Scientific Journal, ESJ, 10*(8).

Ausubel, D. P. (2000). *The Acquisition and Retention of Knowledge: A Cognitive View*. Boston, MA: Kluwer Academic.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1968). *Educational Psychology: A Cognitive View*. New York, NY: Holt, Rinehart and Winston.

Barnes, T., & Stamper, J. (2010). Automatic Hint Generation for Logic Proof Tutoring Using Historical Data. *Journal of Educational Technology and Society, 13*(1), 3-12.

Berlanga, A. J., Van Rosmalen, P., Boshuizen, H. P., & Sloep, P. B. (2012). Exploring Formative Feedback on Textual Assignments with the Help of Automatically Created Visual Representations. *Journal of Computer Assisted Learning, 28*(2), 146-160.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education, 5*(1), 7-74.

Boren, T., & Ramey, J. (2000). Thinking Aloud: Reconciling Theory and Practice. *IEEE Transactions on Professional Communication, 43*(3), 261-278.

Boud, D. (2000). Sustainable Assessment: Rethinking Assessment for the Learning Society. *Studies in Continuing Education, 22*(2), 151-167.

Boud, D., & Molloy, E. (2013). Rethinking Models of Feedback for Learning: the Challenge of Design. *Assessment and Evaluation in Higher Education, 38*(6), 698-712.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience, and School (Expanded Edition)*. Washington, DC: National Research Council.

Cañas, A. J., & Carvalho, M. (2004). *Concept Maps and AI: An Unlikely Marriage.* Paper presented at the SBIE 2004: Simpósio Brasileiro de Informática Educativa, Manaus, Brasil.

Carter, M. (2007). Ways of Knowing, Doing, and Writing in the Disciplines. *College Composition and Communication,* 385-418.

Chatel, R. G. (1997). *Writing to Learn in Science: A Curriculum Guide*. Coventry Science Center, Coventry, CT: Educational Resources Information Center (ERIC).

Chen, N. S., Wei, C. W., & Chen, H. J. (2008). Mining e-Learning Domain Concept Map from Academic Articles. *Computers and Education, 50*(3), 1009-1021.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science, 5*(2), 121-152.

Comer, D. K., Clark, C. R., & Canelas, D. A. (2014). Writing to Learn and Learning to Write Across the Disciplines: Peer-to-Peer Writing in Introductory-Level MOOCs. *International Review of Research in Open and Distance Learning, 15*(5), 26-82.

Croasdell, D. T., Freeman, L. A., & Urbaczewski, A. (2003). Concept Maps for Teaching and Assessment. *Communications of the Association for Information Systems, 12*(1), 24.

Darling-Hammond, L. (2015). *The Flat World and Education: How America's Commitment to Equity Will Determine Our Future*. New York, NY: Teachers College Press.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 319-340.

De Souza, F., Boeres, M., Cury, D., De Menezes, C., & Carlesso, G. (2008). *An Approach to Comparison of Concept Maps Represented by Graphs*. Paper presented at the Third International Conference on Concept Mapping, Tallinn, Estonia and Helsinki, Finland.

Deane Sorcinelli, M., & Elbow, P. (1997). *Writing to Learn: Strategies for Assigning and Responding to Writing Across the Disciplines*. San Francisco, California: Jossey-Bass.

Debuse, J. C., Lawley, M., & Shibl, R. (2008). Educators' Perceptions of Automated Feedback Systems. *Australasian Journal of Educational Technology, 24*(4).

Dikli, S. (2006). An Overview of Automated Scoring of Essays. *The Journal of Technology, Learning and Assessment, 5*(1).

Elbow, P. (1994). Writing for Learning -- Not Just for Demonstrating Learning. Amherst: University of Massachusetts, Amherst, National Teaching and Learning Forum.

Emig, J. (1977). Writing as a Mode of Learning. *College Composition and Communication, 28*(2), 122-128.

Enger, S. K. (1996). *Concept Mapping: Visualizing Student Understanding*. Paper presented at Annual Meeting of the Mid-South Educational Research Association, Tuscaloosa, Alabama.

Ferrara, L., & Butcher, K. R. (2011). *Visualizing Feedback: Using Graphical Cues to Promote Self-Regulated Learning.* Paper presented at the Thirty-third Annual Conference of the Cognitive Science Society, Austin, TX.

Ferster, B., Hammond, T. C., Curby Alexander, R., & Lyman, H. (2012). Automated Formative Assessment as a Tool to Scaffold Student Documentary Writing. *Journal of Interactive Learning Research, 23*(1), 79.

Floyd, R. W. (1962). Algorithm 97: Shortest Path. *Communications of the ACM, 5*(6), 345.

Foltz, P. W., & Rosenstein, M. (2015). *Analysis of a Large-Scale Formative Writing Assessment System with Automated Feedback*. Paper presented at the Second ACM Conference on Learning @ Scale, Vancouver, BC, Canada.

Forsman, S. (1985). Writing to Learn Means Learning to Think. *Roots in the Sawdust*, 162-174.

Glynn, S. M., & Muth, K. D. (1994). Reading and Writing to Learn Science: Achieving Scientific Literacy. *Journal of Research in Science Teaching, 31*(9), 1057-1073.

Gouli, E., Gogoulou, A., Papanikolaou, K., & Grigoriadou, M. (2004). *COMPASS: an Adaptive Web-based Concept Map Assessment Tool*. Paper presented at the First International Conference on Concept Mapping, Pamplona, Spain.

Harlen, W., & James, M. (1997). Assessment and Learning: Differences and Relationships Between Formative and Summative Assessment. *Assessment in Education, 4*(3), 365-379.

Hartmann, D. P. (1977). Considerations in the Choice of Interobserver Reliability Estimates. Journal of Applied Behavior Analysis, 10(1), 103-116.

Hartmann, S., Szarvas, G., & Gurevych, I. (2012). Mining Multiword Terms from Wikipedia. *Semi-Automatic Ontology Development: Processes and Resources*, 226-258.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81-112.

Hay, D., Kinchin, I., & Lygo Baker, S. (2008). Making Learning Visible: the Role of Concept Mapping in Higher Education. *Studies in Higher Education, 33*(3), 295-311.

Hay, D. B. (2007). Using Concept Maps to Measure Deep, Surface and Non-Learning Outcomes. *Studies in Higher Education, 32*(1), 39-57.

Hiebert, J., & Carpenter, T. P. (1992). Learning and Teaching with Understanding. In D. Grouws (Ed.), *Handbook of Research on Mathematics Learning and Teaching* (pp. 65-97). New York, NY: Macmillan Publishing Company.

Hsieh, S.-H., Lin, H.-T., Chi, N.-W., Chou, K.-W., & Lin, K.-Y. (2011). Enabling the Development of Base Domain Ontology through Extraction of Knowledge from Engineering Domain Handbooks. *Advanced Engineering Informatics, 25*(2), 288-296.

Ifenthaler, D. (2010). Bridging the Gap Between Expert-Novice Differences: The Model-based Feedback Approach. *Journal of Research on Technology in Education, 43*(2), 103-117.

Jiang, B., & North-Samardzic, A. (2015). *Acceptance and Use of Moodle by Students and Academics*. Paper presented at the Twenty-first Americas Conference on Information Systems (AMCIS), Puerto Rico, USA.

Jonassen, D. H., Reeves, T. C., Hong, N., Harvey, D., & Peters, K. (1997). Concept Mapping as Cognitive Learning and Assessment Tools. *Journal of Interactive Learning Research, 8*(3), 289.

Kintsch, E., Steinhart, D., Stahl, G., LSA Research Group, L. R. G., Matthews, C., & Lamb, R. (2000). Developing Summarization Skills through the Use of LSA-based Feedback. *Interactive Learning Environments, 8*(2), 87-109.

Kuh, G. D. (2008). Excerpt from High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter. *Association of American Colleges and Universities, 14*(3), 28-29.

Lachner, A., Burkhart, C., & Nückles, M. (2017). Formative Computer-Based Feedback in the University Classroom: Specific Concept Maps Scaffold Students' Writing. *Computers in Human Behavior, 72*(Supplement C), 459-469.

Larranaga, M., Conde, A., Calvo, I., Elorriaga, J. A., & Arruarte, A. (2014). Automatic Generation of the Domain Module from Electronic Textbooks: Method and Validation. *IEEE Transactions on Knowledge and Data Engineering, 26*(1), 69-82.

Leake, A. V. D. (2006). *Jump-Starting Concept Map Construction with Knowledge Extracted from Documents*. Paper presented at the Second International Conference on Concept Mapping, San José, Costa Rica.

Liao, Q., Shim, J. P., & Luo, X. (2004). *Student Acceptance of Web-Based Learning Environment: an Empirical Investigation of an Undergraduate IS Course.* Paper presented at the Tenth Americas Conference on Information Systems, New York, NY.

Lukasenko, R., Anohina-Naumeca, A., Vilkelis, M., & Grundspenkis, J. (2010). Feedback in the Concept Map Based Intelligent Knowledge Assessment System. *Scientific Journal of Riga Technical University. Computer Science. Applied Computer Systems, 41*(1), 17-26.

Madnani, N., & Cahill, A. (2018). *Automated Scoring: Beyond Natural Language Processing.* Paper presented at the 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, USA.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). *The Stanford CoreNLP Natural Language Processing Toolkit*. Paper presented at the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations Baltimore, Maryland.

Marchewka, J. T., & Kostiwa, K. (2007). An Application of the UTAUT Model for Understanding Student Perceptions Using Course Management Software. *Communications of the IIMA, 7*(2), 10.

Mathur, R. (2011). Students' Perceptions of a Mobile Application for College Course Management Systems (Doctoral Dissertation). *Available from ProQuest Dissertations and Theses Database.(UMI No. 3466835)*.

Morton, J. (2007). Authenticity in the IELTS Academic Module Writing Test: a Comparative Study of Task 2 Items and University Assignments. *IELTS Collected Papers: Research in Speaking and Writing Assessment, 1*, 197-248.

Nathawitharana, N., Huang, Q., Ong, K.-L., Vitartas, P., Jayaratne, M., Alahakoon, D., . . . Ahmed, T. (2017). Towards Next Generation Rubrics: an Automated Assignment Feedback System. *Australasian Journal of Information Systems, 21*.

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Committee on the Foundations of Assessment. Pelligrino, J., Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative Assessment and Self-Regulated Learning: a Model and Seven Principles of Good Feedback Practice. *Studies in Higher Education, 31*(2), 199-218.

Novak, J. D. (2002). Meaningful Learning: The Essential Factor for Conceptual Change in Limited or Inappropriate Propositional Hierarchies Leading to Empowerment of Learners. *Science Education, 86*(4), 548-571.

Novak, J. D. (2005). Results and Implications of a 12-Year Longitudinal Study of Science Concept Learning. *Research in Science Education, 35*(1), 23-40.

Novak, J. D. (2010). Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations. *Journal of E-Learning and Knowledge Society, 6*(3), 21-30.

Novak, J. D., & Canas, A. J. (2007). Theoretical Origins of Concept Maps, How to Construct Them and Uses in Education. *Reflecting Education, 3*(1), 29-42.

Novak, J. D., & Cañas, A. J. (2008). *The Theory Underlying Concept Maps and How to Construct and Use them (Technical Report IHMC CmapTools 2006-01 Rev 01-2008)*. Florida Institute for Human and Machine Cognition.

Novak, J. D., & Gowin, D. B. (1984). *Learning How to Learn*. New York, NY: Cambridge University Press.

O'donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge Maps as Scaffolds for Cognitive Processing. *Educational Psychology Review, 14*(1), 71-86.

Olney, A. M., Cade, W. L., & Williams, C. (2011). *Generating Concept Map Exercises from Textbooks.* Paper presented at the 6th Workshop on Innovative Use of NLP for Building Educational Applications, Portland, Oregon.

Oye, N., Iahad, A., Rahim, N., & Zairah, N. (2012). A Comparative Study of Acceptance and Use of ICT Among University Academic Staff of ADSU and LASU: Nigeria. *International Journal of Science and Technology, 1*(1), 40-52.

Özbek, E. A. (2016). Plagiarism Detection Services for Formative Feedback and Assessment: Example of Turnitin. *Journal of Educational and Instructional Studies in the World, 6*(3), 64-72.

Palermo, C., & Thomson, M. M. (2019). Classroom Applications of Automated Writing Evaluation: A Qualitative Examination of Automated Feedback. In *Educational Technology and the New World of Persistent Learning* (pp. 145-175): IGI Global (Idea Group Publishing), Hershey, Pennsylvania.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond Cold Conceptual Change: The Role of Motivational Beliefs and Classroom Contextual Factors in the Process of Conceptual Change. *Review of Educational Research, 63*(2), 167-199.

Reynolds, J. A., Thaiss, C., Katkin, W., & Thompson Jr, R. J. (2012). Writing-to-Learn in Undergraduate Science Education: a Community-Based, Conceptually Driven Approach. *CBE (Cell Biology Education)—Life Sciences Education, 11*(1), 17-25.

Richardson, R., Goertzel, B., Fox, E. A., & Pinto, H. (2006). *Automatic Creation and Translation of Concept Maps for Computer Science-Related Theses and Dissertations.* Paper presented at the Second International Conference on Concept Mapping, San José, Costa Rica.

Rivard, L. O. P. (1994). A Review of Writing to Learn in Science: Implications for Practice and Research. *Journal of Research in Science Teaching, 31*(9), 969-983.

Romance, N. R., & Vitale, M. R. (1999). Concept Mapping as a Tool for Learning: Broadening the Framework for Student-Centered Instruction. *College Teaching, 47*(2), 74-79.

Roscoe, R. D., Allen, L. K., Johnson, A. C., & McNamara, D. S. (2018). *Automated Writing Instruction and Feedback: Instructional Mode, Attitudes, and Revising.* Paper presented at the Human Factors and Ergonomics Society Annual Meeting, Los Angeles, CA.

Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and Procedures for Assessing Cognitive Skills. *Review of Educational Research, 63*(2), 201-243.

Sadler, D. R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science, 18*(2), 119-144.

Sadler, D. R. (1998). Formative Assessment: Revisiting the Territory. *Assessment in Education: Principles, Policy and Practice, 5*(1), 77-84.

Sengul, S., & Senay, S. C. (2014). Assessment of Concept Maps Generated by Undergraduate Students About the Function Concept. *Procedia - Social and Behavioral Sciences, 116*, 729-733.

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research, 78*(1), 153-189.

Stemler, S. E. (2004). A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. *Practical Assessment, Research & Evaluation, 9*(4), 1-19.

Suskie, L. (2018). *Assessing Student Learning: A Common Sense Guide*. San Francisco, CA:Jossey-Bass.

Szajna, B. (1996). Empirical Evaluation of the Revised Technology Acceptance Model. *Management Science, 42*(1), 85-92.

Terzis, V., & Economides, A. A. (2011). The Acceptance and Use of Computer Based Assessment. *Computers and Education, 56*(4), 1032-1044.

Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding Feedback: A Learning Theory Perspective. *Educational Research Review, 9,* 1-15.

Trumpower, D. L., & Sarwar, G. S. (2010). *Formative Structural Assessment: Using Concept Maps as Assessment for Learning.* Paper presented at the Fourth International Conference on Concept Mapping, Viña del Mar, Chile.

U.S. Department of Education, Office of Educational Technology. (2015). *Ed Tech Developer's Guide*. Washington, D.C.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2016). Unified Theory of Acceptance and Use of Technology: a Synthesis and the Road Ahead. *Journal of the Association for Information Systems*, *17*(5), 328-376.

Ventura, J., & Silva, J. (2012). Mining Concepts from Texts. *Procedia Computer Science, 9*, 27-36.

Villalon, J., & Calvo, R. A. (2009). *Concept Extraction from Student Essays, Towards Concept Map Mining.* Paper presented at the Ninth IEEE International Conference on Advanced Learning Technologies, Riga, Latvia.

Villalon, J., & Calvo, R. A. (2011). Concept Maps as Cognitive Visualizations of Writing Assignments. *Educational Technology and Society, 14*(3), 16-27.

Villalon, J., Calvo, R. A., & Montenegro, R. (2010). *Analysis of a Gold Standard for Concept Map Mining – How Humans Summarize Text Using Concept Maps.* Paper presented at the Fourth International Conference on Concept Mapping, Viña del Mar, Chile.

Villalon, J., Kearney, P., Calvo, R., & Reimann, P. (2008). *Glosser: Enhanced Feedback for Student Writing Tasks.* Paper presented at the Eighth IEEE International Conference on Advanced Learning Technologies Santander, Cantabria.

Villalon, J. J., & Calvo, R. A. (2008). *Concept Map Mining: A Definition and a Framework for its Evaluation.* Paper presented at the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, NSW, Australia

Wan Mohamed, W. A., & Omar, B. (2008). *Using Concept Map to Facilitate Writing Assignment*. Paper presented at the Proceeding of the Third International Conference on Concept Mapping, Tallinn, Estonia and Helsinki, Finland.

Wang, T.-I., Su, C.-Y., & Hsieh, T.-C. (2011). Accumulating and Visualising Tacit Knowledge of Teachers on Educational Assessments. *Computers and Education, 57*(4), 2212-2223.

Warschauer, M., & Ware, P. (2006). Automated Writing Evaluation: Defining the Classroom Research Agenda. *Language Teaching Research, 10*(2), 157-180.

Whitelock, D. (2018). Advice for Action with Automatic Feedback Systems. In *Software Data Engineering for Network eLearning Environments* (pp. 139-160): Springer.

Whitelock, D., Twiner, A., Richardson, J. T., Field, D., & Pulman, S. (2015). *OpenEssayist: A Supply and Demand Learning Analytics Tool for Drafting Academic Essays.* Paper presented at the Fifth International Conference on Learning Analytics and Knowledge, Poughkeepsie, NY, USA.

Williams, C. G. (1998). Using Concept Maps to Assess Conceptual Knowledge of Function. *Journal for Research in Mathematics Education, 29*(4), 414-421.

Wilson, J., & Andrada, G. N. (2016). Using Automated Feedback to Improve Writing Quality: Opportunities and Challenges. In *Handbook of Research on Technology Tools for Real-World Skill Development* (pp. 679-704): IGI Global (Idea Group Publishing), Hershey, Pennsylvania.

Wu, D., & Hiltz, S. R. (2004). Predicting Learning from Asynchronous Online Discussions. *Journal of Asynchronous Learning Networks, 8*(2), 139-152.

Zinsser, W. K. (1989). *Writing to Learn.* New York, NY: Harper and Row.

Zubrinic, K., Kalpic, D., & Milicevic, M. (2012). The Automatic Creation of Concept Maps from Documents Written Using Morphologically Rich Languages. *Expert Systems with Applications, 39*(16), 12709-12718.