Examining Interventions and Cognitive Load Factors in Online Learning Experiences

by

Seok-Joo Kwak

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in the University of Michigan
2020

Doctoral Committee:

Assistant Professor Joi-Lynn Mondisa, Chair
Professor Seth Guikema
Associate Professor Cong Shi
Research Professor Stephanie Teasley

Seok-Joo Kwak

seokjook@umich.edu

ORCID iD: 0000-0001-9142-0809

Acknowledgments

This would not have been possible without a tremendous amount of support from mom, dad, Jinny, and Bo. Also, I am extremely grateful to Hai, Sim, and Jang for being amazing friends and providing me with support.

I would also like to extend my deepest gratitude to my advisor, Dr. Joi-Lynn Mondisa. Thank you so much, for being an amazing mentor, and providing guidance throughout my PhD journey.

I would like to express my deepest appreciation to my committee members.

Thank you to Dr. Cong Shi for providing me with research opportunities and advising me and being an inspiring mentor, and role model throughout my studies here at the University of Michigan.

Thank you to Dr. Stephanie Teasley, for providing me with a tremendous amount of help and feedback for this dissertation and helping me to improve my research skills.

Thank you to Dr. Seth Guikema for being an amazing PhD graduate chair advisor, and for being an encouraging and understanding committee member.

I cannot begin to express my thanks to Dr. Elizabeth Hildinger for understanding and helping me with the many edits of my dissertation and for helping me to better express my thoughts and ideas.

I would like to also take this opportunity to thank God and everyone who supported me and helped me along the way.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Since the beginning of the development of massive open online courses (MOOCs), these and other online learning environments have been considered as potential partial solutions to some persistent problems in higher education. These learning environments, while they have great educational value, have not been as effective as they could be, because they have largely been built with little or no foundation in the cognitive processes (e.g., the conversion of items from short-term to long-term memory) required for effective and efficient online learning. Many innovative online learning approaches are in development, such as personalized learning (learning experiences tailored to address particular information that students need) using adaptive learning systems (machine learning techniques used by computers to recommend materials). However, these approaches would also benefit from being grounded in cognitive theory to better reveal how learning occurs in these systems. Furthermore, crucial features of interventions in online learning, such as supplementary elements designed to fill in gaps or reinforce knowledge, have not been thoroughly examined in conjunction with the insights of cognitive theory and the concept of desirable difficulty (i.e., the notion that the addition of difficulty to a task can improve learning and increase retention).

In this exploratory work, I experimentally examine five different types of interventions and their effects on undergraduate engineering students' learning gains and experience. This study presents quantitative research along with detailed qualitative thematic analysis. Its objective is to provide critical insights into how to better design online learning environments and how we can create more effective interventions that promote students' online learning

gains. The research questions for this work are: (1) What factors in online learning environments affect *learning gains* (i.e., measured difference between post- and pre-test scores) for undergraduate engineering students?; (2) What factors in online learning environments affect the *learning experience* for undergraduate engineering students, and, specifically, what factors produce desirable difficulty?; and (3) What factors in online learning affect undergraduate engineering students' *self-reported memory*?

The experimental results, examined within the framework of cognitive theory, showed quantitatively that levels of frustration with interventions were correlated with learning gains while qualitative analysis results revealed instances that both confirmed and contradicted aspects of the quantitative results. A number of practical design guidelines emerged from the analysis: for example, in specific circumstances, one type of intervention is likely to be more effective than another, or that particular sorts of additional difficulties should be avoided. These recommendations may provide researchers with a better understanding of how to challenge students in more efficient and productive ways in online learning environments.

Chapter 1 Introduction

Distance learning, a phenomenon in which students are presented with educational material remotely via technology-based devices, has taken on a significant role in the world of education (Casey, 2008; Jeffries, 2009). This should come as no surprise: every time a new technology has entered society, educational researchers have been quick to put effort into making use of it in education. In the early 1900s, technology-based distance education started with the introduction of audio devices (e.g., radio) into the schools (Casey, 2008; Jeffries, 2009; Schlosser, 1996). During the middle of the 20th century, television-based education began serving as a means of delivering education for educational institutions and the military (Casey, 2008; Jeffries, 2009; Schlosser, 1996). Starting from the early 1990s, educational researchers have experimented with online-based distance learning (i.e., online learning), especially in higher education (Casey, 2008; Jeffries, 2009). The introduction of the World Wide Web in 1989 changed the landscape of research in the delivery of education; many educators and researchers began realizing that the internet might make it possible to reach students anywhere in the world and provide educational information quickly in a variety of formats.

Research shows that there are tremendous potential benefits to online learning (Gardner & Brooks, 2018). In the early development of online learning, the primary work was focused on how to make it most effective for students (e.g., improving the aesthetic design of online learning, finding an optimal length for videos) (Guo, Kim, & Rubin, 2014; Pomales-Garcia & Liu, 2006). With the rise of massive open online courses (MOOCs), YouTube, and other online educational platforms, millions of students were provided access to high-quality learning resources at their convenience and at little or no cost (Kaplan & Haenlein, 2016). In recent decades, the focus has been shifting to include how to respond to particular students and to better tailor learning to each individual student. Many innovative approaches are in development, but at the same time, educational researchers have also been investigating problems with higher education that have become increasingly pressing.

1.1 Problems with Higher Education

Even as we enter an age supported by state-of-the-art technologies, there still exist formidable problems with the quality of education worldwide (Malcom-Piqueux & Bensimon, 2017). Some of the most persistent include (1) rising college tuition costs; (2) disparities in the educational and social backgrounds of students; and (3) a lack of quality in higher education due to high student-faculty ratios. In the paragraphs below, I discuss each of these problems in detail.

1.1.1 Rising College Tuition Costs

Paying for higher education has become significantly financially stressful for college students across the country and is a source of anxiety for recent graduates burdened with massive student debt (Grabmeier, 2015). In the early 1980s, a student could afford to pay a full year's tuition at a four-year university by working a minimum-wage job. According to the U.S. Bureau of Labor Statistics, from 1983 to 2017, the Consumer Price Index (CPI), an index that measures changes in the price level of a market basket of consumer goods and services purchased by households, increased by a percentage ranging from 100% to 246% for all items. However, college tuition increased by 833% which is three times more than the CPI increase of all items. Furthermore, 68% of bachelor's degree recipients graduated with student loan debt at an average of $30,100 per borrower (DiGangi, 2017). Additionally, 40% of student borrowers are not making student loan payments to the federal bank, which in most cases provided the loans

(Mitchell, 2016), and the burden of these loans shifts to the taxpayer. Students may also have a hard time making the transition to the adult world, because they cannot save and thus may not be able to purchase homes or automobiles, etc. Furthermore, they may feel that their career choices are constrained; a student who wants to work in a rewarding but relatively low-paying field like social work may be reluctant to do so, since his or her low earnings may make repayment of loan debt an intolerable financial burden. Therefore, the rising cost of education is a severe concern for students who will incur college debt and should be for taxpayers as well because of the societal costs.

1.1.2 Educational and Social Disparities

Generally, higher education is viewed as an opportunity that brings social mobility and equality, but the few selected top-tier universities historically may have perpetuated stratification rather than weakening it (Freedman, 2013), because the upper and upper-middle classes dominate access to top private universities. In first world nations, top-tier universities are well funded and have a clear mission and well-designed academic programs (Jacob, Xiong, & Ye, 2015). Furthermore, those top-tier institutions consist of high-quality faculty, well-prepared students, and sufficient resources. However, in developing countries or low-income cities in the United States, most higher education institutions suffer severe deficiencies in each of these areas (Jacob et al., 2015). It is reasonable to assume that, in these same areas, secondary education is also frequently deficient, and students from low-performing or ill-funded schools may graduate with significant gaps or deficiencies that render them inadequately prepared for higher education. Consequently, the education gap between the rich and the poor grows wider. It is likely that everyone would consider it a societal good if all students, regardless of their income and educational preparedness, could receive the benefit of and enjoy the same access to high-quality faculty from top-tier universities at a manageable cost.

1.1.3 Lack of Quality Education Due to High Student-Faculty Ratios

In addition to problems of quality resulting from the conditions noted in the previous section, an issue in higher education is that large class sizes may detract from students' overall experiences. For example, one measurement index of the quality of higher education institutions is the student-faculty ratio (i.e., the number of students divided by the number of faculty at that institution). Obviously, higher enrollment equates to higher ratios. Increasing enrollment has become more common as universities undertake cost control measures, but low student-faculty ratios are preferred for educational reasons (Centra, 2009). It seems self-evident that the lower ratios give students more opportunities to ask questions during lectures and to build networks with other students and faculty members. On the other hand, from universities' business perspectives, a high student-faculty ratio allows the school to generate more student tuition per faculty salary. Hence, the university presumably shares the educational goals but is constrained by financial realities to consider financial objectives as well.

1.2 Potential Solutions for Higher Education

While online learning was never envisioned as the solution to all of these problems, it has the potential to ameliorate some of them because it provides:

- ***Free and Low-cost Access to MOOCs.*** Selective universities and companies have been developing and offering free or low-cost MOOCs for the public (Baturay, 2015); this enables students who could not afford to attend these universities to benefit from the education they provide.

- ***Access to Educational Content from Top Tier Institutions' Courses.*** Online learning can provide its users with access to educational content from prestigious institutions as well as the flexibility to access courses asynchronously (Marek et al., 2015).
- ***Learner-centered Pacing.*** MOOCs offer students the opportunity to take a course at a convenient time and at their own pace (Y. Zhang, 2013); thus they can provide quality education to people at various socio-economic and educational levels and in various contexts to help reduce educational disparities.

These features and the flexible, usually modular, structure of online learning environments--they may include multimedia modules, tests/quizzes and online forums, for example--enable online learning to be a contributor to a solution; more details will be discussed in the literature review.

However, for online learning to reach its full potential, four critical challenges have to be met. The first challenge is that, because online learning is open to everyone who has internet access, it can be perceived to be less valuable than traditional university instruction even when it is offered by prestigious universities (Keramida, 2015). A second problem that is commonly mentioned is the low course completion rate for online learning, particularly in MOOCs, despite their convenience as to their pace and their availability (Hew & Cheung, 2014). A third problem, particularly in earlier forms of online learning, has been students' insufficient prerequisite knowledge. If students have insufficient prerequisite knowledge about the topic being presented, they may have a limited understanding of the material and their progress through the course may come to a halt. A fourth problem has been that bad experiences due to poor delivery of information in an online learning environment can discourage students from continuing and completing courses (Onah, Sinclair, & Boyatt, 2014). Therefore, researchers have been developing ways to meet these challenges. For this work, I have concentrated on the third and fourth challenges listed above; the first and the second (i.e., perceptions of being less valuable and low completion rates) are outside the scope of this research.

In this study, I focus on ways of addressing students' lack of knowledge and problems of online learning environment design. Both of these aspects have been the subjects of studies in recent years. First, researchers have been actively studying the role of interventions (supplementary elements designed to fill in gaps or reinforce knowledge) in addressing students' lack of knowledge and exploring how these can be most effectively delivered. Researchers and educators have also been looking at creating personalized online learning, that is, online learning experiences in which students receive information that fills the gap in their knowledge and is tailored to the particular things that they do not understand. One of the ways to create personalized online learning is to make use of adaptive learning systems or intelligent tutoring systems. This is an approach that involves incorporating machine learning techniques, in which computers build a mathematical model based on sample data of previous students to recommend appropriately tailored videos, text, or tasks to help students to understand the material that has been or will be presented to them.

While personalized learning is becoming increasingly important in our efforts to provide better learning experiences, it is crucial that before focusing specifically on personalization, we ensure that our approaches to designs of learning are rooted in a firm understanding of good and effective ways to deliver information in general, how the conditions of students act as factors in learning, and how different pedagogies make a difference in learning outcomes. O. Chen, Woolcott, and Sweller (2017) suggest that MOOCs should be grounded in an understanding of the cognitive processes required for effective and efficient online learning. While Sweller's most recent work concerns MOOCs specifically, the points he makes can probably be applied to broader questions; most current MOOCs are similar to other online learning environments (that is, they tend to include adaptive elements that aim to provide immediate and customized instruction or feedback to learners via quizzes and interventions). Current research on personalized online learning also suggests that behavioral patterns such as boredom or frustration in online learning tasks may pose problems that may be found also in other forms of online learning and that should be explored as well. In fact, frustration is one of the most commonly mentioned negative emotions in studies of online learning in general (Capdeferro & Romero, 2012) and frustration is one of the key reasons for learners' high dropout rates in MOOCs (Capdeferro & Romero, 2012), though some research also shows that it can potentially help students to become motivated (Radel, Pelletier, Baxter, Fournier, & Sarrazin, 2014). All this research itself suggests that MOOCs, as well as other online learning environments, should be rooted in fundamentals of delivering materials efficiently and also focus on understanding learners' behaviors to help them learn the best. Once we have a solid understanding of the fundamentals, then we can narrowly focus our attention on personalized learning or on any new forms of educational designs that may arise.

1.3 The Purposes of This Study

The purposes of this study are (1) to provide an in-depth exploration of factors that affect students' learning experience in an online learning environment; (2) to illuminate the features of interventions that affect undergraduate engineering students' online learning experience; and (3) to investigate relationships between factors in order to demonstrate mutual influences, both positive and negative.

The value of this research is that it will provide insights into better online learning environment design and ways of creating more effective learning materials, especially interventions, that promote students' online learning gains and improve their learning experience.

1.4 Overview of the Methodology

This dissertation consists of a pilot study and an exploratory experimental study to examine the online learning experiences of engineering undergraduates. In the initial pilot study reported in Chapter 3, I investigated the experience of undergraduate engineering students who were given a video intervention during their engagement in an online learning task. The objective was to identify their perceptions of an adaptive learning environment that used MOOCs materials (in this case, videos from different courses). To characterize their perceptions and the effect of the intervention on their learning experience, I collected survey data, interview data, and post-test scores for 18 students (in a simulated adaptive learning environment). After collecting the data, I analyzed them using basic statistical methods for the quantitative data and thematic analysis for the qualitative data. Two key results emerged from the pilot study:

(1) Students seem to have found the adaptive learning experience enjoyable even though their post-test scores were low.

(2) Students' frustration with the adaptive learning tasks may be linked to the monotony of the video instructor or the students' own lack of content knowledge.

From these results, I created two informal hypotheses. First, I hypothesized that there may be a negative correlation between students' learning gains and their perceptions of enjoying a learning experience including an adaptive task. This finding suggested that the perception of adaptive learning needs to be further investigated and that using an existing cognitive learning framework would be appropriate. Thus, I decided to use cognitive load theory, a theory that captures the way that learners process information in short- and long-term memory, as a sensitizing framework for my dissertation research. Cognitive load theory captures the way information is processed in memory. The theory also touches on the optimal use of auditory (hearing) and visual (seeing) channels for processing information.

Second, I hypothesized that perhaps students received the intervention information too easily from the video and that there may exist other types of interventions that may create better learning scores by requiring students to put in more effort. This hypothesis led me to discover a body of educational research focusing on the important concept of creating desirable difficulty. According to Bjork (1994), the term *desirable difficulty* refers to the concept that additional difficulty imposed on a learning task for the purpose of increasing recall, retention, and long-term learning gains. It should be noted that “desirable” implies limits; too much difficulty is undesirable and produces non-value-added frustration or boredom. Both the pilot study findings and the literature following Bjork’s work (see Section 2.7) suggested a need for further research on which types of interventions and which uses of the learners' channels (i.e., auditory and visual) in these interventions can cause desirable difficulties that may result in better learning gains and improve self-reported memory.

As previously noted, both quantitative and qualitative methods were used for the pilot study. This approach is not common in research that examines online learning. Typically, many quantitative studies examine the effects of personalized online learning in which additional material is recommended to students. Also, there is an abundance of research, primarily quantitative, on best practices in online learning in general. While these have much to offer in regard to the questions I wished to investigate, the qualitative piece that is not well represented in the literature is important for getting at the best ways to address the goals of maximizing learning, because it captures students' perspective in ways that quantitative methods cannot easily do (Creswell, 2002). While there may be general studies that look at how students' learning gains and self-reported memory are affected by interventions using the two learning channels (e.g., auditory and visual), to my knowledge, currently there is a lack of research that examines these outcomes in combination, focusing on the ways and varying extents to which the channels are used in the interventions. This dissertation brings together elements from these earlier works and examines them in relation to each other and to the original work presented here. Thus, it makes a contribution by synthesizing these to produce new insights and to answer some questions perhaps left unanswered by the earlier work.

In the exploratory experimental study, I examine five different types of interventions (in the forms of Audio-only, Text-only, Video, Video+Text, and a writing task) in an online learning environment and capture how they affect undergraduate engineering students' learning gains and self-reported memory of content presented to them. Because this experiment was designed to simulate cases in online learning environments where the students receive the intervention, each student was provided with an intervention regardless of whether he/she would have received it in a real case (which would have been determined by his/her performance on an assessment instrument). In this study, the independent variables are demographic information, intervention type, delivery of information, and pedagogical approach; the dependent variables are learning gains, learning experience, and self-reported memory. The investigation of the research questions proceeds by means of quantitative research in conjunction with qualitative thematic analysis. The three central research questions addressed in this dissertation are:

1. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning environments affect *learning gains* (i.e., measured difference between the post-test and the pre-test scores) for undergraduate engineering students?
2. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning environments affect the *learning experience* for undergraduate engineering students, and, specifically, what factors produce desirable difficulty?
3. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning affect undergraduate engineering students' *self-reported memory*?

In order to answer these questions, I conducted an experiment to test my hypotheses and explored the results. This dissertation is organized as follows: in Chapter Two, I discuss the literature relevant to my research. In Chapter Three, I report on the pilot study I conducted to gather preliminary data to inform the research design of my dissertation study. In Chapter Four, I explain the methods I used to conduct this research. In Chapter Five, I report the results of this research. In Chapter Six, I discuss the results and findings. And finally, in Chapter Seven, I present my conclusions and recommendations for future work.

Chapter 2 Literature Review

In recent years, a great deal of research has been directed at questions of how the increasing use of electronic media is helping students learn better. In an effort to ensure the best learning outcome and experiences for students, researchers have been investigating various techniques, approaches, and designs to improve online education. In the following sections, I focus on the major themes that have garnered a great deal of attention in this literature. First, I briefly revisit some basic questions: what online (multimedia) learning is, what its uses in higher education contexts are at present, what benefits it offers, and what its drawbacks or shortcomings are. Next, I specifically discuss two major problems of online learning (lack of prerequisite knowledge and poor environment designs) and their corresponding potential solutions (further development of personalized learning and application of cognitive load theory). Finally, I conclude with a discussion of desirable difficulty, a factor in online learning related to cognitive load theory.

2.1 What Is Online Learning and Who Uses It?

As defined by prominent researcher Dr. Richard Mayer, multimedia learning is "learning that involves learning from words and pictures and includes learning from textbooks that contain text and illustrations, computer-based lessons that contain animation and narration, and face-to-face slide presentations that contain graphics and spoken words" (Mayer, 2014). The use of multimedia learning has increased greatly as Internet use has increased (Mast, 2015). Multimedia learning using the internet is what we now call an online learning environment. Online learning is an instructional mode in which students are presented with course information remotely via online media (Pomales-Garcia & Liu, 2006). In online education, the teaching media currently exist in many forms (e.g., video courses, blogs, podcasts, e-mails, instant messages, chat rooms, and forums) (Pomales-Garcia & Liu, 2006). Online learning platforms (e.g., MOOCs, Khan Academy, Udacity, Coursera, etc.) use various combinations of these forms to create a course.

Online learning is used by many different types of people for a wide variety of purposes. Since 2015, K-12 education enrollment in the United States has reached 3.8 million (Zheng, Lin, & Kwon, 2020). Online learning for K-12 is used mainly to make up for a shortage of courses for remedial or accelerated students and a lack of access to qualified teachers in local schools (Cavanaugh & Clark, 2007). For higher education, the uses are more commonly the replacement or supplementation of classroom instruction: as of 2015, approximately a quarter of all college students (6 million) in the United States were taking an online class. Another group making use of online learning is post-graduates; according to a survey conducted in 2014, 84 percent of online students identified as working professionals whose purposes for enrolling in online classes were job-related, personal interest/lifelong learning, or interest in the MOOC format (Liu et al., 2014). Online education can be used in both formal classrooms and informal learning spaces, and since learning materials can be accessed from laptops, smartphones, and tablets, it is available essentially everywhere a person can connect to the Internet (Gutiérrez-Rojas, Alario-Hoyos, Pérez-Sanagustín, Leony, & Delgado-Kloos, 2014). The primary focus of this research is examining online usage in higher education contexts, where it is playing an increasingly important role in broadening access to high-quality education throughout the world.

2.2 Online Learning and Its Uses in Higher Education

As briefly indicated above, research studies show that online learning is currently being used in various ways in higher education, such as formal instruction, instructional support, and informal self-education. The uses have been evolving for many years, both as the technologies for delivery and consumption of online learning products have developed and as our understanding of how learning in the online environment occurs. In 2008, George Siemens and Stephen Downes, two educational researchers, pursued their goal of finding out whether it was feasible to teach a massive number of participants (Downes, 2008). They led an open online course that they made available for a fee to 25 paying students and free to an additional 2300 students (Herman, 2012). Dave Cormier at the University of Prince Edward Island and Bryan Alexander of the National Institute for Technology in Liberal Education, who were working in the area, coined the term massive open online courses, or MOOCs to describe this type of course (Leito, Helm, & Jalukse, 2015). Siemens and Downes' overall goal was to use the Internet to reach a massive number of participants (Downes, 2008). Since 2008, many MOOCs platforms have been created (e.g., edX, Udacity, Coursera, OpenLearning, Class2Go, 10genEducation, Khan Academy, etc.) (Stevens, 2013).

Recently, significant numbers of students with diverse backgrounds have begun to make use of online learning. Allen and Seaman (2017) report that 6 million undergraduate students are taking at least one course online; thus, we can assume that significant use of online learning may be affecting university students. In a descriptive and experimental study on college students (n = 91), Jaffar (2012) demonstrated that 98% of university medical students were using online learning videos as a source of information. It is also noteworthy that 81 of the top 100 universities ranked by Times Higher Education World University Rankings in 2015 offered MOOCs (Shigeta et al., 2017).

Because the use of online learning seems to be a growing trend, researchers are exploring how university students are affected by their increasingly frequent experiences with online learning environments (Iniesto, McAndrew, Minocha, & Coughlan, 2016). Much of the research quantitatively investigates how online learning can help students, but quantitative methods are limited in their ability to capture and explain learners' experiences and perspectives (Veletsianos, 2013). There is a lack of qualitative research on online learning. In the research reported in this dissertation, I combine the use of quantitative and qualitative methods in an attempt to fill in the gap about students' online learning experiences. Before we look at the research questions in detail, however, it will be useful to identify many of the positive and negative features of online learning; these will be discussed in the following paragraphs.

2.3 Positive Features of Online Learning

Currently, in higher education, there are educational gaps, social disparities, and problems resulting from lesser quality education (Malcom-Piqueux & Bensimon, 2017). Many studies suggest that online learning can address these issues because it has features that are well suited to reducing costs, improving access to high-quality education, and ameliorating inequities. In this section, several positive features of online learning that address these issues will be discussed.

2.3.1 Financial Benefits for Students and Institutions

As mentioned briefly in the introduction, online learning has financial benefits for both the students and the universities. Since the introduction of MOOCs, selective universities and companies have been developing and offering free or low-cost MOOCs for the public (Baturay, 2015); this enables students who could not afford to attend these universities to benefit from the education they provide. For traditional university education, it is hard to know what the exact cost for a single class is, but, according to Xing and Marwala (2017), the two biggest factors that impact the cost of development of a class in traditional universities are physical costs (e.g., building costs, maintenance costs, equipment costs, etc.) and productivity costs (e.g., the salaries of faculty members, salaries of support staff persons, cost of insurance, etc.). According to

Kirkham (2018), the typical cost of one college credit for one student comes out to $594 if averaged across every sector (including private and public, for-profit and not-for-profit, and two- and four-year colleges). This implies that a 3-credit course costs $1,782 ($594x3 = $1,782). This high cost per student is typically reflected in tuition cost per credit hour, which makes traditional university education very costly for students. In contrast, for an equivalent MOOCs educational offering, the cost of development has been shown to range from about $28,980 to $325,330; the cost per student completing the course is about $74 to $272 (Hollands & Tirthali, 2014). Thus, for a fraction of the cost or no cost, any students, whether or not they are enrolled, can have access to high-quality education provided by MOOCs, as top-tier universities are giving more public access to their courses.

The return on investment (ROI) for the university is extraordinarily high with MOOCs. From the university's business perspective, MOOCs are financially advantageous because they have an extremely high student-to-teacher ratio, which reduces the cost per student (Wu, Daskalakis, Kaashoek, Tzamos, & Weinberg, 2015). Also, MOOCs simplify scheduling and logistics for universities; they can be accessed and used by learners at any time, never require breaks from lectures, and reduce or eliminate many physical and personnel costs (i.e., insurance costs, benefits, etc.). Clearly, online learning has financial advantages for both students and institutions.

2.3.2 Great Lectures from Prestigious Institutions

Online learning can provide its users with the flexibility to access courses asynchronously as well as access to educational content from prestigious institutions. Many of the universities participating in partnerships with MOOCs providers, such as Coursera, are listed among the top universities in the world, such as Harvard, Stanford, Princeton, MIT, University of Pennsylvania, University of California - Berkeley, the University of Michigan, etc. (Marek et al., 2015). In 2020, Coursera alone offered some 3000 courses, many of them taught for students by prominent faculty members from these top-tier universities (Abbakumov, Desmet, & Van den Noortgate, 2020). This potentially addresses the educational gap by giving students who would not otherwise have a chance to attend these schools the opportunity to have great lectures from great universities.

2.3.3 Flexibility and Accessibility to Information

MOOCs offer students the opportunity to take a course at a convenient time and to move through it at their own pace (Y. Zhang, 2013), thus they can provide quality education to people at various socioeconomic and educational levels and in various contexts to help reduce educational disparities. As of 2020, the largest MOOCs provider in the world, Coursera, has reached 36 million students (Abbakumov et al., 2020).

These online courses can be accessed from any location without admission to the top universities, as stated above (Walia, 2020). This online learning can reach employed people who want to enrich their knowledge without meeting the academic qualifications necessary for enrolling in regular courses or committing themselves to an academic program (Walia, 2020). Another big benefit is that learners do not have to be in physical classrooms; they can be in different parts of the world and be able to access high-quality education from wherever they are (Singh, 2020). The flexibility of online courses allows self-pacing at students' own convenience (Jansen & Schuwer, 2015). Online learning gives learners flexibility in choosing topics that they want to learn because of the low cost of online learning (Castillo, Lee, Zahra, & Wagner, 2015). This flexibility of online learning also benefits teachers, who can spend time developing new classes or conducting research if they are free from the need to continually develop traditional course offerings. Once an online course is created, it can be repeatedly accessed, and no further effort is required of the creator.

2.4 Negative Features of Online Learning

Despite the many benefits of online learning, many negative features and disadvantages prevent online courses and environments from achieving all the goals of their designers. A number of problems, weaknesses, and challenges, several of them interrelated, can be identified; several of these are discussed below.

2.4.1 Lack of Certification and Non-Acceptance for College

The first challenge is that, even though online education options like MOOCs may be offered by prestigious universities, because they are open to everyone who has internet access, they are perceived to be less valuable than traditional university instruction (Keramida, 2015). Currently, very few universities accept certifications of online course completion (e.g. badges, MOOCs certificates, certification of specialized training) as adequate to justify awarding course credit for their completion (Singh, 2020). Furthermore, it is not clear how online learning certificates are valued in the labor market (Singh, 2020). For example, employers may view online learning as less rigorous (no lab work, homework, tests) than traditional university offerings. One can argue that students are receiving education of equal value in online courses but, as recognized in standard marketing principles, because this education is free, learners may not perceive it as valuable. From the student's perspective, there are also other benefits to traditional class instruction, such as networking with professors and peers and hands-on experiences, that online classes simply cannot offer.

The challenge here is how to bring about the integration of online learning into college curricula so that students might realize the cost savings. That will require that designers of online courses find ways to achieve the level of rigor of traditional classes, to maintain quality control, and to compensate in some ways for the absence of hands-on course components. Thus, one important goal of online learning proponents is finding a way to change the perception of online learning and integrate online learning classes into university curricula (Walia, 2020). If universities treat online courses (MOOCs) as fully equivalent to in-person course offerings, it is likely that employers will also recognize their value, which will make it more feasible for students to make use of this low-cost option in the pursuit of a college degree (Walia, 2020).

2.4.2 High Dropout Rate

The second disadvantage of online learning education that is commonly mentioned is the low course completion rate for online learning, particularly in MOOCs, even though they offer convenience as to their pacing and their availability (Hew & Cheung, 2014). Currently, according to Rout, Sahoo, and Das (2020), the completion rate for those who register for online courses, in general, is only about 5 - 10%. In the field, a MOOC is considered 'completed' if a learner actually watches the first module and continues watching to the point that he/she reaches the final task or presentation in the course (Bárcena, Read, Martín-Monje, & Castrillo, 2014). However, before we proceed to explore the real problem with completion rates, it is essential to understand that this very low number somewhat misrepresents the real situation because there are many factors contributing to the high drop-out rate that do not reflect a problem or are even positive.

2.4.2.1 Learners' Supplemental or Non-Academic Uses of OLE

Two contributors to the high drop-out rate that are not necessarily negatives are (1) students' use of online courses as a supplement to their formal education and (2) students' enrolling just to pursue personal interest (i.e., just to check the material out).

Research shows that some persons who enroll in online courses are using the courses merely as a supplement to the instruction that they are receiving in their classes or on the job (Cutrell et al., 2015; Onah et al., 2014). A number of studies have found that the vast majority of online learners already have a college degree and thus may not be using online courses toward a degree (Christensen et al., 2013; Despujol, Turró, Busquéis, & Cañero, 2014). It seems reasonable to speculate that these users may be using online learning environments (OLE) for work-related purposes or personal interest; in such cases, failure to complete a course may mean that the user achieved his or her purpose for enrolling the course at some point before its end.

Another factor that contributes to the high dropout rate that is not necessarily negative is that students may just be curious about a topic. Since many courses are free, they can sign up for the course just to check it out and, in the absence of any financial commitment, they may feel no pressure to complete it (Bárcena et al., 2014).

These reasons do not suggest any deficiency in the online resources they use; presumably, many learners benefit even if they do not complete the whole course (Parr, 2012). For the reasons articulated above, these cases should not be included in the drop-out rate but should rather be understood as different cases that actually achieved their purpose of helping students.

2.4.2.2 Reasons for Dropping out that are Relevant & Addressable

There are two significant reasons for low completion rates that are relevant and can be addressed in the design of an online learning environment: (1) students' inability to grasp the material and their resulting frustration, and (2) bad learning experiences resulting from poor instructional design. While there may be additional valid reasons for students to drop out, these are, to the best of my knowledge, the primary problems that designers of OLE can potentially solve, and thus they will be examined further below.

2.4.3 Students' Lack of Knowledge May Lead to Dropping Out

One problem is that the level of difficulty of a course may be too great for students who lack sufficient prerequisite knowledge. If there is no adequate support for these learners to address the issue, then this leads to an overall lack of understanding of the topic, which may make students drop out because of frustration (Onah et al., 2014). In a study conducted at Duke University, many students were not able to complete an online course (bio-engineering) specifically because of difficulty with the mathematical requirements, and this topic arose frequently on the discussion boards (an online site for questions and comments about the course) as well. Insufficient prior content knowledge is a major obstacle for students in completing online courses; (Belanger & Thornton, 2013) observed that the problem was exacerbated because students had nowhere to turn to address their insufficient knowledge. Also, even students with adequate prerequisite knowledge, if they cannot ask questions about new materials they are learning, may become frustrated by the lack of real-time feedback in their online course, and this may lead them to drop out (Khalil & Ebner, 2014). Thus, it is clear that the absence of a means for compensating a lack of knowledge and addressing students' questions as they are learning can lead to their failure to complete online courses.

2.4.4 Problems with Online Learning Environment (OLE) Designs

The second problem is that bad experiences in an online learning environment can discourage students from continuing; (Onah et al., 2014) speculate that bad online learning experiences result from poorly designed learning environments that do not convey information effectively. Currently, most MOOCs' design is based on learners' opinions and feedback about the quality of the instruction, even though learners generally do not have the expertise to evaluate the educational design (Marginson, 2016). According to O. Chen et al. (2017), since the beginning of the development of MOOCs, these have largely been built with little or no foundation in the cognitive processes required for effective and efficient online learning. It has been recognized that reducing the learner's mental effort is important for effectiveness and efficiency, but that depends on the careful design of the learning environment, and that has not received the attention that it needs. A study conducted by Marginson (2016) indicated that the course evaluation of online learning education typically does not take instrumental design quality (e.g., delivery of the education materials, interactiveness, presenters' quality) as part of their evaluations.

Researchers investigating these aspects have particularly identified several questions that they believe should be posed, such as: Are learners receiving correct and appropriate feedback (i.e., is there adequate interactivity)?, Are learning materials logically structured and easy to find for students (i.e., are students unlikely to be frustrated by difficulties unrelated to the actual content)?, Are various learning forms (Text, Audio, Video) provided for students (i.e., are students likely to find materials that are best adapted to their learning preferences)?, etc. Although some MOOCs have been designed with attention to ways of maintaining student engagement --for example, engagement (how the length of a presentation of some content is related to learners' engagement), rates of learner engagement and persistence remain low (Sari, Bonk, & Zhu, 2020). To the best of my knowledge, there is not much research on how cognitive processes are applied to adaptive learning environments to improve effectiveness.

2.5 Ways to Address Online Learning Issues

Despite all of the aforementioned potential problems with online learning, there are a few pertinent solutions to some of these problems. As I mentioned above, problems related to the perceived value of online learning and the high dropout rate, which results in part from supplemental and curiosity-driven uses, are irrelevant and beyond the scope of this research. However, existing literature may provide suggestions for addressing other problems related to online learning. Specifically, students' lack of knowledge and poor design may be addressed using adaptive learning and a fundamental understanding of the application of cognitive load theory. In the following sections, I review the literature on adaptive learning systems and the application of cognitive load theory to designing environments.

2.5.1 Personalized Learning and the Use of Adaptive Learning Systems

To address the problem of insufficient prerequisite knowledge or a limited grasp of new knowledge in an online learning environment, some researchers have investigated the development of personalized learning (Brown, 2015; Li, Xu, Zhang, & Chang, 2020). The term *personalized learning* refers to approaches designed to give students what they need at a given moment on the basis of the individual student's current knowledge or behavior, to maximize the learning objective either through human intervention or, in automated systems, by employing an algorithm. Although this study does not directly focus on the implementation of personalized learning, it was undertaken with that as a goal for future work. Thus, it is important to understand what personalized learning is and what its attributes are.

2.5.2 How Personalized Learning Addresses Students' Lack of Knowledge

To create personalized online learning, researchers have been investigating adaptive learning systems or intelligent tutoring systems. While these two terms are now commonly used as near-synonyms, adaptive learning systems are in a sense an outgrowth of early research into intelligent tutoring systems (Hatzilygeroudis & Prentzas, 2009; Phobun & Vicheanpanya, 2010; Weber, 2012). In the pre-internet years, development of intelligent tutoring systems was limited because of the difficulty of building large databases that would be suitable for each single topic (Hatzilygeroudis & Prentzas, 2009). However, with the explosion of data on students' learning behavior and the increasing sophistication of machine learning techniques in recent years, they have become an important focus of research in the personalization of online instruction (Phobun & Vicheanpanya, 2010). Both intelligent tutoring systems and adaptive learning systems use algorithms that can analyze learners' attributes (i.e., current knowledge or behavior states) to find personalized learning paths and thus to choose the most optimal learning materials or path (Li et al., 2020; Xie, Chu, Hwang, & Wang, 2019).

In the following section, I provide a few examples of current adaptive learning environments. Many quantitative data analyses (e.g., machine learning techniques, Markov models, factor analysis) have been used in exploring designs for personalized online learning environments. For example, in 2016, Williams et al. (2016) developed a system called AXIS (Adaptive eXplanation Improvement System) that asked learners to generate an explanation for another learner and used machine learning to evaluate and identify the best explanations among the learners' explanations for a future student. The system showed some initial promise in generating optimal explanations that help students. Researchers have used a hidden Markov model to track students' progress and to find personalized learning paths for students (Y. Chen, Culpepper, Wang, & Douglas, 2018). Other researchers have used deep learning techniques that help students memorize material more effectively (Reddy, Levine, & Dragan, 2017). In most cases, algorithms make the judgment as to what students need and recommend appropriate supplementary learning materials, which might take a variety of forms of intervention (e.g., videos, text, audio, etc.).

As suggested above, there exists a vast literature using quantitative methods, while qualitative insights and perceptions of adaptive tasks in personalized learning remain largely unexplored. For example, Rosen et al. (2018) suggest that students' behavioral patterns, such as boredom or frustration in adaptive tasks, have not been adequately treated and should be explored in order to improve adaptive learning environments. Liu et al. (2014) used sensor-free observation methods to qualitatively study how confusion and frustration may affect students' online learning where these are associated positively with learning outcomes for short materials and negatively for lengthy materials. As researchers are increasingly applying adaptive learning techniques to the online environment, the learner's perspective is becoming increasingly important to consider in the design of this environment.

2.6 Using Cognitive Load Theory to Address OLE Designs

In this section, I will explain the basics of cognitive load theory and explain cognitive loads in greater detail.

2.6.1 What is Cognitive Load Theory?

To address the problems of online learning design, since the early 2000s, researchers have been applying cognitive load theory to the design of multimedia learning materials to help increase the effectiveness of the materials (Brame, 2016). Cognitive load theory draws on the Atkinson-Shiffrin model of memory, which posits that memory has three components: (1) sensory memory, (2) working memory, and (3) long-term memory (Atkinson & Shiffrin, 1968). The model posits that memory is stored through the process laid out below:

1. When a learner tries to learn something, words or pictures are presented to and received by the learner's verbal/auditory channel or visual/pictorial channel.
2. Information from the two channels is collected in ***sensory memory***.
3. From sensory memory, the information goes into ***working memory*** (also known as short-term memory), where it is processed and organized. This component is extremely limited in both how much it can hold and how long it can hold it.
4. From the working memory, a few selected pieces of information are moved to ***long-term memory*** and stored. This component is unlimited in both how much it can hold and how long it can hold it. From the long-term memory, information can be retrieved by working memory when needed. This process is captured in the diagram in Figure 2-a below, which has been simplified from Atkinson and Shiffrin (1968).

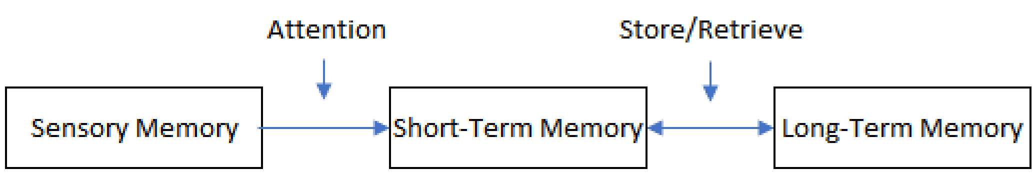


Figure 2-a: The Atkinson-Shiffrin Model of Memory Simplified

Sweller's insight was that because working memory has limited capacity and only a small quantity of information can be selected for storage in the long term memory, designers of learning materials must figure out how to avoid overloading the working memory with non-value-added information (Sweller, 1988, 1989, 1994). Thus, Sweller developed cognitive load theory, which categorizes the information that a learner's mental capacity must be handled as three types of "loads": (1) intrinsic cognitive load (ICL), (2) extraneous cognitive load (ECL), and (3) germane cognitive load (GCL).

2.6.2 Intrinsic Cognitive Load, Extraneous Cognitive Load, and Germane Cognitive Load

Intrinsic cognitive load is the mental effort demanded by the inherent complexity of the learning task and material (Cooper, 1998). For example, the level of difficulty of quantum physics has a relatively high inherent difficulty; the level of difficulty of a simple addition task (e.g., 1+1=2) has relatively low inherent difficulty. *Extraneous cognitive load* is cognitive work that does not have a payoff in learning; for example, distractions such as background noise in lecture videos or unreadable handwriting increase the difficulty of the learning task but have not been shown to increase learning (De Jong, 2010). *Germane cognitive load* is the cognitive work necessary to connect incoming information to existing knowledge; the term also refers to the product of the work. For example, when a learner is able to understand a new concept such as multiplication through associating it with the familiar context of addition, the learner is making the new concept germane and assimilating it to existing knowledge.

Exploring applications of Atkinson and Shiffrin's model in conjunction with cognitive load theory, Mayer and Moreno (2003) proposed a cognitive theory of multimedia learning based on the dual-channel assumption and limited-capacity assumption. Their theory states that a limited capacity of the learner’s brain selects and obtains a multimedia presentation (e.g., text, pictures, and auditory information) through the learner’s dual-channel and selects and organizes the presentation dynamically to produce logical mental constructs rather than interpreting them mutually exclusively.

Positing limitations on the cognitive load that can be processed and the high selectivity of the brain, Mayer, Moreno, and many other cognitivist learning theorists have studied ways to create effective learning designs by setting the objective of a learning outcome to minimize extraneous cognitive load, increase germane cognitive load, and manage intrinsic cognitive load

(Brame, 2016). They found that a wide variety of ways to achieve this objective had been developed in recent years, including segmentation, signaling, matching modalities, and weeding, all of which I discuss in the following paragraphs.

Segmentation is simply the division of a body of information into small sections to ease the management of intrinsic load and increase the germane load. Studies have shown that maintaining students' attention is difficult after 13 minutes and that presenting material in 3-6 minute videos leads to engagement of students, as measured by their willingness to continue watching up to 100% of the time (Guo et al., 2014). When learners are engaged, they can process the incoming information so that it becomes germane. D. Zhang, Zhou, Briggs, and Nunamaker Jr (2006) examined the influence of interactive videos on learning outcomes and learner satisfaction in online learning environments. Their results showed that segmenting a video is critical for students' engagement. Ibrahim, Antonenko, Greenwood, and Wheeler (2012) conducted a study showing that segmenting a video into smaller units enabled students to transfer knowledge better and reported lower levels of learning difficulty.

Signaling is a strategy used to reduce extraneous load. Mayer and Johnson (2008) conducted experiments on college students in which they had two groups: one group's lecture slides contained 2–3 signaling words (i.e., short, redundant words) that were identical to the words in the lectures' speech and the other group's did not. Results showed that the students whose presentation included short redundant words outperformed the non-redundant group on a subsequent test of retention, on the basis of which they concluded that signaling and redundancy reduced extraneous load. Moreno and Mayer (2007) analyzed the effect of directing attention to relevant information with signaling and segmentation in dynamic instructional videos by creating one with signaling and segmentation and one without. The findings showed that, while the control group outperformed the signaling and segmentation group on the retention of theoretical information, the signaling and segmentation group performed better when asked to evaluate what they learned and to apply teaching skills in a classroom scenario. The signaling and segmentation group appears to have had lower levels of cognitive load. Within the framework of cognitive load theory, signaling is understood to reduce time spent identifying key ideas in lecture slides and thus to reduce extraneous cognitive load and enhance learning.

Matching modality is a strategy predicated on the assumption that learners can better manage the cognitive load if the proper channels are activated (Mayer & Moreno, 2003). Research suggests that using audiovisual materials selectively to activate visual and auditory channels appropriately would increase student engagement with videos to provide flexibility in learning experiences (Thomson, Bridgstock, & Willems, 2014). One additional common practice is weeding (i.e., reducing background noise or eliminating extra animation that does not add value). Weeding minimizes the extraneous load so that more of the learners' cognitive capacity can be used for the germane load; (Ibrahim et al., 2012). Mayer and Johnson (2008) has also conducted research exploring the redundancy effect; the study found that it tends to show reduced extraneous load processing. These applications of cognitive load theory to multimedia learning have spurred numerous advances designed to ease the cognitive load placed on students.

A typical MOOC lesson demonstrates how these concepts underlie the design: it takes approximately 30 minutes and the lesson is composed of 4-9 minute of modules, tests and quizzes, and various tasks (Abbakumov et al., 2020). The features and components named above reflect the research-informed choices that the online learning designers must make and take into consideration how the length, difficulty, order, etc. of the materials affect students' learning in various ways. These have been based primarily on quantitative analysis. However, there is a lack of research looking at quantitative and qualitative data in combination. Furthermore, at the same time that the research was showing positive effects of cognitive load minimizing strategies, a seemingly contradictory idea called desirable difficulty began to be investigated, in which making things harder rather than easier is posited to have long-term benefits.

2.7 Desirable Difficulty as a Factor in Online Learning

Since 1994, when Robert A. Bjork coined the term *desirable difficulty* to refer to additional difficulty imposed on a learning task for the purpose of increasing recall, retention, and long-term learning gains (Bjork, 1994), many researchers have explored this concept in many experiments. Soderstrom and Bjork (2015) provide a comprehensive review of difficulty-inducing techniques that have clear practical benefits for long-term learning but may negatively affect short-term learning. Two of several well-known effective desirable difficulty techniques are retrieval practice (i.e., activities such as flashcards or testing that are concurrent with the presentation of the material to be learned), and spacing effects (i.e., spreading out practice sessions over a period of time). Both of these techniques force the learner to retrieve the information from memory, in one case immediately and in the other at several times after the content to be learned has been presented.

As an example of retrieval practice, Marsh and Butler (2013) have shown that when students use flashcards that require them to answer questions rather than simply re-reading their class notes, their recall of information appeared to improve. As an example of spacing practice, researchers have also been testing whether delaying feedback to students rather than giving immediate feedback leads to better learning outcomes, but the results show that the effect of delayed feedback is not well understood (Swanwick, 2013). While retrieval practice is a well-established desirable difficulty practice, there are still some open questions that need to be addressed, such as how to sequence test questions so as to maximize long-term memory and what type of retrieval practice condition produces the best benefits (Heitmann, Grund, Berthold, Fries, & Roelle, 2018).

Varying the conditions of practice is another technique that has been studied. This refers to adding difficulty by varying the conditions (or context) of learning rather than keeping them constant and predictable. For example, researchers have experimented with learners studying the same material in two different rooms rather than twice in the same room (Smith, Glenberg, & Bjork, 1978). The study reported that changing the condition of the learning environment leads to increased recall of that material. However, in other studies, Paas and Van Merriënboer (1994) tested the effect on learning of adding difficulty by having high variability in the format of questions and comparing it with low variability; the result did not indicate that high-variability conditions were always more effective in increasing learning than low variability. Another study testing the effect of varying the conditions of practice (O. Chen, Castro-Alonso, Paas, & Sweller, 2018) showed that this may not always produce desirable difficulty.

As mentioned above, researchers have been attempting to ascertain whether adding difficulty for students consistently leads to better learning outcomes. Even among researchers who accept the notion that desirable difficulty aids learning, there seems to be some contradiction as to what constitutes desirable difficulty. O. Chen et al. (2018) has argued that if there is not enough working memory capacity to deal with the new information or set of tasks, then many difficulties we add to information will be undesirable, as might occur when capacity is within working memory limits (O. Chen et al., 2018). The varying and sometimes conflicting conclusions indicate that more studies need to be done.

2.8 Writing Reflections in Online Learning Environments

Historically, the term reflection in learning has referred to the process of examining one's knowledge and learning (Dewey, 1933; Peltier, Hay, & Drago, 2005; Rogers, 2001). The writing reflection, especially, has also been studied as an activity that might be beneficial for student learning. According to Cowan (2014), the task of writing a reflection helps learners to think as they attempt to answer a question(s) from the point of view that is practical for them. Alsanad, Howard, and Williamson (2016) state that the task of writing a reflection provides the learners an opportunity to analyze and synthesize information from their point of view.

According to Jansen and Schuwer (2015), writing about what one is learning takes in multiple processes: (1) learners comprehend the lecture material, (2) they identify key points, (3) they link the material to their prior knowledge and prior notes, (4) they paraphrase or summarize, and (5) they transform the material to written form (either by hand or by typing). While this study focused on note-taking as the writing task, its findings may illustrate why a writing reflection is helpful from a cognitive load theory perspective, in that learners who do such self-reflecting processes have to apply their cognitive resources to these processes about the topic they are learning. Applying the desirable difficulty framework to writing tasks, Suzuki, Nakata, and Dekeyser (2019) state that writing tasks have been known to reduce the cognitive processing that helps with students' knowledge. In such cases, long-term learning gains may not occur.

In MOOCs or online learning, writing reflection takes various forms and serves various purposes. One representative example of how a writing reflection is used in online learning or MOOCs is a design in which, while students are learning about a topic, there are written assignments and blog posts in which they are asked primarily to reflect on the concepts but also to comment on their learning processes O'Brien, Forte, Mackey, and Jacobson (2017). Another example of the use of a writing reflection is given in a study by Williams et al. (2016), in which the system asked a learner to provide written self-explanations for other learners. Since instructors have limited time and resources to generate quality explanations, researchers created a learning platform that prompts learners to write explanations for a topic they are learning. After each learner explains, the system iteratively refines the explanation (i.e., the intended outcome, via combining all the explanations provided and using machine learning (on the set of explanations) to choose the most effective elements of the submissions). Through this process, the most helpful explanation is constructed for the learners.

These examples show that a writing reflection is perceived to have value for designers of MOOCs and other similar learning environments, and that there is general agreement that it allows students to engage deeply and interactively with the topic they are learning. However, the practice has not been extensively studied in combination with some other well-researched aspects in online learning, such as cognitive load theory, desirable difficulty, and behavior patterns, particularly frustration and boredom, all of which contribute significantly to long-term learning.

2.9 Frustration in Online Learning Environments

According to Capdeferro and Romero (2012), frustration is one of the most commonly mentioned negative emotions in studies of online learning. Frustration has been defined in several ways, but the definitions are all similar. According to Mandler (1975), frustration is a negative emotional response aroused upon encountering an obstacle in the achievement of a task, goal, or expectation.

According to Iepsen, Bercht, and Reategui (2013), frustration plays a significant role in the experience of students and also affects their cognitive processes in learning. They conducted research to search for patterns in the learners' quantitative behavior data and qualitative data that indicate frustration when learners are working in a multimedia learning environment. When learners showed behaviors associated with frustration in their learning processes, researchers were able to help learners by focusing on and helping them overcome their difficulties by providing them with resources.

There are several ways to measure frustration. It has been measured with survey questions incorporating Likert scales (e.g., measuring the degree to which a person can tolerate frustration or feels frustrated in a situation using a 4-, 5-, or 7-point Likert scale (Harrington, 2005; Peters, O'Connor, & Rudolf, 1980; Wright, Lam, & Brown, 2009). In other studies, learners' heart rate and facial expressions were measured using photoplethysmogram signal sensors and cameras to implicitly infer their emotional state in MOOCs (boredom, confusion, curiosity, frustration, happiness, and self-efficacy) (Pham & Wang, 2017). With these data collected, researchers were able to detect and understand learners' moment-to-moment emotion states, which mean that incorporating an understanding of emotions into the design of online learning environments could potentially improve outcomes (Xiao, Pham, & Wang, 2017). In 2017, novice students (n=99) participated in a self-paced computerized learning environment experiment intended to detect and identify affective states during learning (Bosch & D'Mello, 2017). Students engaged in a learning task and their facial expressions were captured at intervals; afterward, they were asked to judge their emotions from the photographs taken during the session. The results indicated that engagement, confusion, frustration, boredom, and curiosity were the most frequent affective states, while anxiety, happiness, anger, surprise, disgust, sadness, and fear were rare. Confusion + frustration and curiosity + engagement were identified as two frequently co-occurring pairs of affective states.

In online education, a poor experience due to slowness, lack of access to needed information, or the presence of unrelated information can cause students to feel frustrated or even to abandon their learning tasks (Conrad, 2002). However, other studies suggest that previously-experienced or socially accepted frustration in learning may help learners develop the motivation to do a new task (Radel et al., 2014) or can even make them try harder try harder in a subsequent learning task (DeWall, Baumeister, & Vohs, 2008). These findings reinforce the idea that there is an important distinction between positive and negative frustration.

To sum up this literature review, on the basis of the research surveyed here, it is clear that there are still questions of interest to people who are studying how to help students learn, particularly in online environments. According to Mayer (2014), cognitive theories of multimedia learning tend to focus on instructional methods aimed at reducing extraneous load or managing intrinsic load, whereas motivational theories tend to focus on instructional methods aimed at challenging the students. This would reinforce the idea that there are two predominant approaches in education: one that posits the value of reducing difficulty (through means such as intelligent tutoring, designing to reduce cognitive load, etc.) and one that favors increasing difficulty (through the application of value-adding frustration). This second approach is related to desirable difficulty, which has been shown to be positive, but it is also associated with some negative outcomes because difficulty could presumably increase the likelihood that students would experience non-value-adding frustration. As in the case of cognitive load theory and desirable difficulty, much of the research on this topic has been quantitative. Qualitative research might shed light on this question by revealing learners' subjective feelings about frustration and different types of imposed difficulties; more mixed-method research must be done before we can arrive at any conclusions. Furthermore, there is a lack of research examining the role of desirable difficulty in adaptive learning systems. A preliminary investigation of these questions was undertaken in a pilot study preceding the main work of this dissertation; the results are reported in Chapter 3. The chapters that follow investigate these questions further.

Chapter 3 A Preliminary Pilot Study[1]

Abstract—Open-access online courses, called massive open online courses (MOOCs), have received much attention from higher education institutions and course designers for their potential to reshape learning opportunities. Among the challenges in learning from MOOCs or in an online setting is that if students have insufficient prerequisite knowledge about the topic being presented, they have a limited understanding of the material and they cannot ask questions in person to clarify their understanding. To address this problem, researchers have been developing adaptive learning technologies. Adaptive learning is a form of learning in which a computer changes the lecture content to best fit a given student based on the student's interactions with the interface. However, current literature suggests that behavioral patterns such as boredom or frustration in adaptive online learning tasks should be explored in order to improve students' learning experiences. This study investigated engineering undergraduate students' perceptions of an adaptive learning environment using MOOCs materials. In this exploratory mixed-methods study, I collected and analyzed survey and interview data and post-test scores for 18 students in our experiment. The results of the analysis suggest a negative correlation in the relationship between students' learning gains and their perceptions of their enjoyment of the videos that they were shown in the MOOC.

Index Terms—Adaptive Learning, MOOCs, Personalized E-Learning

3.1 Introduction

Even as we enter an age supported by state-of-the-art technologies, there still exist formidable inequities (i.e., educational gaps, social disparities, lesser quality education) in the quality of higher education worldwide (Malcom-Piqueux & Bensimon, 2017). One possible way

[1] Submitted manuscript: Kwak, S.J. & Mondisa, J.L., Exploring Engineering Undergraduates' Frustration and Enjoyment in Adaptive E-Learning Activities: A Mixed-Methods Study.

to address these inequities is to assess ways to effectively use the content of massive open online courses (MOOCs), which are open-access online courses that permit unlimited participation (Kaplan & Haenlein, 2016). One benefit of MOOCs is that they can provide high-quality learning resources for millions of students to access at their convenience and at no cost (Kaplan & Haenlein, 2016). However, four concerns with MOOCs are: (1) low retention and completion rates; (2) insufficient prior knowledge (3) the lack of help in cases when the learner fails to understand the material; and (4) the lack of interaction between instructors and students (M. Zhang, Zhu, Wang, & Chen, 2018). To address these problems, adaptive learning systems have been the subject of many types of research for the last five years (Rosen et al., 2018). As noted above, adaptive learning systems are those in which a computer changes the online lecture content to best fit a student based on the student's interactions with the interface. The computer uses a machine-learning algorithm incorporating the student's data (e.g., demographic data, GPA, zip code, quiz scores) to recommend what the student should see next. With each student's record and information about what other students were given in similar situations, the machine learning algorithm seeks to optimize each student's learning experience. Some existing research on adaptive learning systems that use machine learning techniques focused on personalized learning provides quantitative results about learning benefits (Beck & Woolf, 2000; Rosen et al., 2018; Williams et al., 2016). Qualitative insights and perceptions of adaptive learning from college students' perspectives, however, are largely lacking. For example, Rosen et al. (2018) provided quantitative evidence on the effects of adaptive learning systems in MOOCs on learning gains; however, the research lacks a qualitative approach and does not illuminate into the underlying reasons, opinions, and motivations as to why students behaved as they did. Furthermore, as mentioned above, Rosen et al. (2018) suggest that behavioral patterns such as boredom or frustration with adaptive learning systems have been insufficiently explored. Therefore, more mixed-method research needs to be done to inform the design of materials that may improve students' learning experiences.

In this preliminary study, I explore the effects of adaptive learning environments from students' perspectives by examining their learning enjoyment, frustration, and use of online learning with e-learning activities. Our project addresses the need to better personalize student e-learning activities by providing detailed qualitative reasoning to help explain quantitative results.

The purpose of this experimental pilot study is to explore how students' adaptive learning experiences influence their levels of frustration and enjoyment in using online learning modules. The main goal of this study was to obtain preliminary findings about the effectiveness of using MOOCs materials and adaptive learning tools. In this study, I examine how students are affected when performing adaptive learning tasks in the context of watching and interacting in an online video lecture setting. The research question guiding this study is: What are students' experiences when engaging in tasks in an adaptive learning environment? Specifically, elements such as learning enjoyment, frustration, and use of online learning during adaptive tasks were explored and analyzed.

In the following sections, I provide background details about the current state of and problems associated with MOOCs, effective video learning, and adaptive learning environments (Section 3.2). Next, I provide details about the experimental design of the study and the methods used to conduct this research (Section 3.3). Section 3.4 presents the quantitative and qualitative results of the study. Next, results of the study are discussed in Section 3.5. Finally, conclusions are outlined in Section 3.6.

3.2 Background

3.2.1 Massive Open Online Courses (MOOCS)

Over the past decade, massive open online courses (MOOCs) have received a great deal of attention in the education field (Gaebel, 2013). MOOCs provide high-quality learning resources for millions of students to access at their convenience, at little or no cost. However, MOOCs come with many challenges. One challenge is that although many students enroll in MOOCs, the retention rates for these courses are very low and only a very small proportion of students complete the courses (Khalil & Ebner, 2014). According to Belanger and Thornton (2013), another challenge is that students who participate in MOOCs may have insufficient prior knowledge about the course topic. This may lead to their becoming frustrated while watching the MOOC and being unable to process the material they are learning. As a result, they may be unable to go on to the next steps in the learning process. Furthermore, while using MOOC content, students may have no one to turn to for help (Belanger & Thornton, 2013). Therefore, another problem with MOOCs is that personalized support is unavailable to students and there is a lack of interaction between instructors and students (M. Zhang et al., 2018).

3.2.2 Effective Video Learning

According to Brame (2016), there are three elements that must be considered in educational video design and implementation in order to keep students engaged and for the video to serve as a productive part of a learning experience. They are: (1) cognitive load (the load that have to do with where the processing of the incoming information takes place); (2) non-cognitive elements that impact engagement (e.g., shortness of content-delivery segments and a conversational style of delivery); and (3) features that promote active learning (e.g., interactive activities, homework).

Also, Guo et al. (2014) state, in order to optimize the cognitive load and keep students engaged during an e-learning experience, it is recommended that videos be kept short, preferably under 6 minutes. Furthermore, appropriately using both auditory and visual channels in videos has been shown to maximize students' retention of the material and increase student engagement (Guo et al., 2014). Guo et al. (2014) also reported that student engagement was dependent on the narrator's speaking rate, such that student engagement increased as the speaking rate increased.

3.2.3 Adaptive Learning Environment

One way to help students and to enhance their online learning experiences would be to exploit intelligent tutoring systems that provide additional explanations of materials to learners (Aleven & Koedinger, 2002). Building on decades of research in intelligent tutoring systems, psychometrics, cognitive learning theory, and data science, researchers have developed adaptive learning systems (Rosen et al., 2018). The defining feature of adaptive learning systems is that a computer algorithm analyzes the student's interactions with the interface and changes the lecture content to best fit that student. However, Rosen et al. (2018) suggested that effectively identifying needed adaptive tasks depends on a better understanding of that behavioral patterns in learners, such as boredom or frustration, in performing adaptive tasks. Thus, these should be explored in order to identify ways to improve adaptive tasks (Rosen et al., 2018). My project is intended to explore the effects of such patterns on students' perceptions of adaptive interactive tasks from students' perspectives and identify potential ways to improve the adaptive learning experience.

3.3 Description of Experiment

The primary purpose of the research study was to explore how students' perceptions of adaptive learning environments are related to frustration with and enjoyment of the modules. To this end, I developed and executed an exploratory mixed methods research design (Creswell, 2002), which involved surveying and interviewing participants to investigate their adaptive learning experiences. The experiment was conducted to test the following hypothesis: As students are exposed to an adaptive learning environment, they may experience frustration but also an increased sense of enjoyment.

I analyzed the survey and interview data to elicit the most emergent themes. Ultimately, I hope to (1) understand whether students become frustrated in engaging in adaptive learning environments and, if they do, why and (2) determine whether students' enjoyment increases as a result of engaging in an adaptive activity. In the following paragraphs, I detail the procedures used to conduct the research.

3.3.1 Recruitment and Selection of Participants

In order to have a population with the same amount of knowledge, it was essential to recruit students with little exposure to the lecture topic featured in the online learning material. This allowed us to measure learning gains across students with the same level of knowledge about the topic. Thus, I initially attempted to recruit University of Michigan (UM) engineering undergraduate students who had not taken any industrial and operations engineering (IOE) courses. My assumption was that if they had not taken any IOE courses, measuring the gain in their learning about the topic would be likelier to be possible In a previous study, Pomales-Garcia and Liu (2006) recruited 18 participants to analyze learners' perceptions and the impact of web modules on their learning experiences. For this pilot study, after receiving approval from the UM Institutional Review Board, I recruited 18 UM engineering undergraduate students via email. I contacted five UM engineering department administrators in the following departments:

mechanical, industrial, biomedical, chemical, and civil. I requested that they distribute the recruitment email through their undergraduate email listservs.

The average age of the sample was 20.38 years and the average GPA was 3.48. As shown in Table 3-1, the majority of participants were from the mechanical engineering and industrial engineering departments. Among the participants, 10 students were male, and 8 students were female. Additionally, the majority of the participants were Asian and White students; see Table 3-2.

Table 3-1: Learner Demographics by Major

Major	Number of Students
Mechanical Engineering (ME)	9
Industrial Engineering (IOE)	5
Chemical Engineering (ChE)	1
Biomedical Engineering (BME)	1
Civil Engineering (CEE)	1
Not yet declared	1

Table 3-2: Learner Demographics by Race/Ethnicity

Race	Number of Students
Asian	9
White/Non-Hispanic	5
Hispanic or Latino	2
American Indian or Alaska Native	1
Black or African American	1

3.3.2 Instruments and Data Collection Procedures

A survey and interview protocol were developed to collect information about participants' experiences in performing adaptive learning tasks in the online learning environment that I designed. Demographic information (i.e., academic major, race, gender, etc.) about participants was also collected.

3.3.2.1 Learning Modules Materials

To examine the effects of the adaptive learning environment, I designed the experiment in three parts: (1) the participants watched a 20- to 30-minute lecture that included adaptive tasks (occurred for every concept that was covered); (2) they were given intermittent assessment (approximately 3 to 10 minutes) of the participants' content knowledge via survey; and (3 they were interviewed by the researcher about their experiences. Specifically, I created an adaptive learning task experience for participants using online learning videos from YouTube and electronic survey software. The topic chosen for the video lecture was basic Economic Order Quantity (EOQ), which focuses on optimizing order quantity to minimize total costs. This concept requires that a student be knowledgeable about economic concepts, calculus, and inventory management. The recruitment method made it likely that the participants would not have been exposed to the EOQ topic and thus would have to learn it from the modules.

I first created the video module by splicing together content from existing YouTube videos about EOQ, from which I selected the ones that had the most views. I then created seven multiple choice questions using Qualtrics software that asked participants questions after each topic was taught.

3.3.2.2 Development of Survey Protocol

The study's survey, consisting of nine questions, was administered immediately after the completion of the learning task. Six questions requested demographic information (e.g., race/ethnicity, sex/gender, academic status, age, major, and citizenship status). The three additional survey items were modified and adapted from Pomales-Garcia and Liu (2006). These questions asked participants to rate their perceptions of the knowledge (i.e., understanding of the material presented in the video) that they had gained using a five-point Likert scale (1 =

completely new material and 5 = expert). This was done to collect information about their perceived adaptive learning experiences in the form of quantifiable data. There was no time limit for answering these questions. These questions were:

a) Before watching the video modules, how much did you know about the topic discussed in the module using a scale of 1-5, where 1 = completely new material and 5 = expert?

b) After watching the video modules, how much did you know about the topic discussed in the module using a scale of 1-5, where 1 = completely new material and 5 = expert?

c) If the rating for the level of difficulty of a children's story for a four-year-old represents a rating of 1, what is the level of difficulty of the content that this module presented?

3.3.2.3 Development of Interview Protocol

I created an interview protocol composed of 17 questions. The interview protocol was developed in collaboration with mixed methods study experts in Dr. John W. Creswell's mixed-method workshop in 2018. The interview questions required the participants to provide more details about their adaptive learning experiences in the study. The first question asked participants about the overall adaptive learning experience. The interview questions focused on three themes: frustration, attention level, and enjoyment of the material, in accordance with the suggestion of Rosen et al. (2018) that behavioral patterns such as boredom or frustration in adaptive tasks should be explored. Specifically, I created interview questions that examined three topics: (1) enjoyment: the enjoyment students experienced in an adaptive online learning environment; (2) frustration: the frustration students experienced in an adaptive online learning environment; and (3) online video usage: students' use of online videos as a supplement to their classroom education. For each topic element, there were corresponding interview questions. For enjoyment, there were six questions (e.g., "What did you like about your experience in completing the module?”). There were six items that measured frustration (e.g., “Describe a time in this process in which you felt frustrated”). For the topic of online video usage, there were two questions (e.g., “Do you use online videos or MOOCs as a supplement in your studying?”). The two final questions inquired whether participants wanted to provide any additional thoughts about the overall experience and, finally, whether there were any final thoughts they would like to add in general.

3.3.2.4 Administering the Survey and Interviews

The experiment was conducted in a closed interview room. The room was equipped with a laptop on which the video module was displayed. The participant station consisted of a Dell laptop running Windows 7 with a screen resolution of 1024*768 pixels at 32 bits of color, with a mouse. I was in the room only to assist the participant with any necessary troubleshooting and to conduct the interview after the student completed the video module.

Survey response data was collected using an online Qualtrics survey. To maintain participants’ privacy, the names of the participants were changed to pseudonyms and any identifiable information was subsequently removed from the reported data. The interview voice recordings were transcribed verbatim by me. Voice recordings were deleted immediately after the transcription. Prior to my conducting this research, this study was approved by the UM Institutional Review Board.

3.3.2.5 Experimental Procedure

The experimental procedure consisted of four steps. First, the researcher explained the outline of the procedure to the participant, and then the participant signed an informed consent form. Next, the participant watched the adaptive video lecture lesson and then took a post-test. The post-test consisted of questions about EOQ topics to assess students’ knowledge of what they had learned in the video. Next, participants completed a survey that collected demographic information and examined their learning experience and, finally, they participated in an interview with me in which data was collected and recorded using the Samsung Galaxy s6 voice recording program. Figure 3-a shows the different steps of the procedure. These procedural steps are discussed in detail in the following sections.

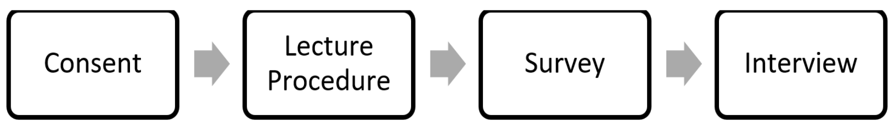


Figure 3-a: General Experiment Procedure

3.3.2.5.1 Lecture Procedure

After consenting, the participants went through a lecture procedure that involved watching a video lecture and taking content quizzes at specific places in the video. Specifically, the participants performed the following steps:

(1) Video: Participants watched a video about an EOQ topic.

(2) Content Quiz: After each topic was explained, the system displayed a short multiple-choice quiz about the topic, as shown in Figure 3-b. Three questions were asked per adaptive task. The adaptive tasks were administered two times. There was one content quiz administered at the end of the lecture procedure.

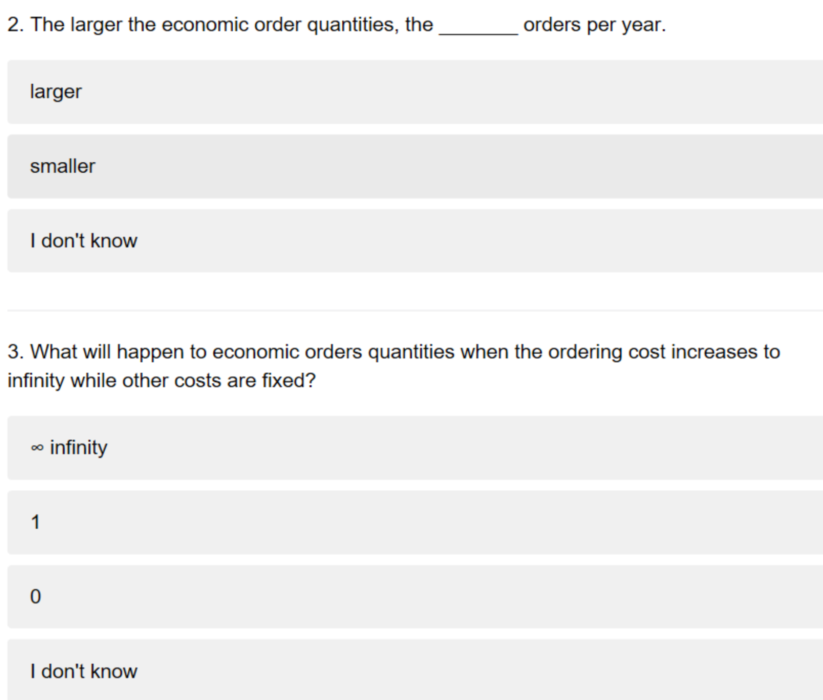


Figure 3-b: An Example of an Adaptive Task Question

(3) Remediation Videos: After the participants completed the content quiz, they were administered one of two types of remediation videos (i.e., short and long). Remediation videos are additional videos that provide a more detailed explanation of the video that they watched. Shorter remediation videos (3 minutes) were shown to the first nine participants who participated in the experiment, and longer remediation videos (8 minutes) were shown to the remaining nine participants. This was done to initially examine how the length of the remediation video affects students. However, I did not find any significant effect on students.

(4) Post-Test: Following the lecture procedure, all participants took a post-test about what they had learned. The post-test consisted of eight questions about EOQ topics to assess students' knowledge of what they had learned during the whole learning process. An example of a post-test question is: "What would happen to economic order quantity if other items

remained the same in the EOQ model, with double annual demand and double the unit cost of purchased materials?"

3.3.2.5.2 Administering the Survey and Interview

After participants completed the post-test, the researcher administered a survey to collect information about participants' demographic characteristics and their experiences. After the participants completed the survey, I conducted interviews with them. The interviews were conducted in a soundproof room. The interviewer asked questions and the interviewee answered in a conversational style. The interview allowed us to learn in detail what the students had experienced during the learning process.

3.4 Data Analysis

After the data collection procedure, I used quantitative (e.g., descriptive statistics) and qualitative (e.g., thematic analysis) methods to analyze the survey and interview data, respectively. Three themes (i.e., Enjoyment, Frustration, Use of Online Learning) were explored qualitatively. First, I performed descriptive statistics analysis and organized students' responses according to the corresponding theme (see Section 3.4.2). I then transcribed the 18 interviews. Finally, I thematically analyzed the interview data to identify major emergent themes (Creswell, 2002). Detailed explanations of the steps performed in analyzing the data quantitatively and qualitatively are provided in the following sections.

3.4.1 Quantitative Phase

I wanted to explore the differences in frustration between students who did extremely well (i.e. scored 100%) on the post-test and students who did not (anything below 100%). Therefore, I divided students into two groups: those with perfect scores and those with nonperfect scores. Then I looked at each score group's transcribed data. On the basis of the transcribed data, I divided the students into two groups who expressed that they were frustrated and those who did not express that they were frustrated about the adaptive learning environment. Next, I tallied how many students were in each category. The findings are displayed in Section 3.5. Within the group of students who expressed frustration, I qualitatively explored why they were frustrated, as will be explained in the next section.

To investigate whether and why students use online videos as educational supplements, I coded the transcribed data into themes. I tallied how many reported that they used online videos to supplement their college study and learning experiences. Then I organized these results, which I present in Section 3.5.

To assess the correlation between students' post-test scores and their perceptions about the enjoyment experienced in the adaptive learning activity, I assessed Pass/Fail scores. First, I collected test score data. Then, a 'Pass' was assigned to scores greater than 70 out of 100 points. Under the enjoyment data, I filtered data corresponding to students who passed and those who failed. Next, I tallied how many students were in each group.

3.4.2 Qualitative Phase

Interview transcripts (n = 18) were thematically coded and used in conjunction with the descriptive statistics information to explore and understand how learners' perceptions of learning environments were affected by the adaptive tasks and the explanation videos. Specifically, thematic analysis (Boyatzis, 1998) was performed using the following method. First, using the interview data, I assessed and categorized the most frequent and common responses about frustration, enjoyment, and the use of online educational videos that arose in the transcripts. Answers that commonly appeared were grouped into the same category theme. For example, six students expressed that they were frustrated with the professor's tone and energy. Their specific responses were grouped into one qualitative category (i.e., frustration) to assist with further interpretation of the quantitative data. Specifically, the interview data allowed us to examine similarities or differences in the interviews of participants and their descriptive statistical data in frustration and enjoyment and the use of videos. The data analysis findings and a discussion of the implications of this research are presented in the next section.

3.5 Results and Discussion

From the data, three main themes emerged about engineering undergraduates' adaptive learning experiences in regard to enjoyment, frustration, and online video usage. Those themes are: (1) Adaptive Learning Environments are Enjoyable, (2) Frustration can be Linked to Teacher Energy and Lack of Student Knowledge, and (3) Students Use Online Video as Supplements to Classroom Education at High Rates.

3.5.1 Adaptive Learning Environments are Enjoyable

My findings seem to suggest that students found the adaptive learning experience enjoyable. It is interesting to note, however, that among all of the students who expressed that the overall adaptive learning process was helpful and enjoyable (n = 15), only approximately half (n = 8) earned a passing quiz score (i.e., more than 70 out of 100 points; see Figure 3-c).

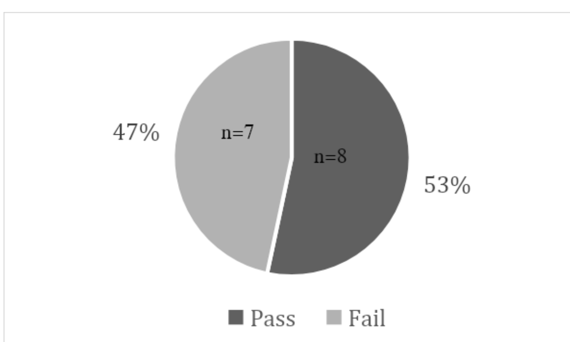


Figure 3-c: Students' Responses to Finds Adaptive Tasks Helpful and Enjoyable (n = 15)

Some of the reasons for professing enjoyment given by students who did not earn a passing score were:

> *"It (the learning experience) was interesting. It (if I already knew the material) would save me a lot of time because if I already learned most of it or just the little things here, then I can move on to different videos."*

> *"It (the learning experience) was good. I liked (that) the quiz asked about content that's included in (the process). That's great because I have that question in mind so the next video, I could expect the video to talk about it."*

> *"I thought it (the learning experience) was great. I feel like this can be used to bridge the gap instead of watching the whole lecture."*

From the data, it seems that students perform more poorly in retaining online content knowledge even though they enjoyed the online video. Specifically, some findings seem to suggest that there is a negative relationship between students' final post-test scores and their perceptions of enjoying the video. This might occur because the videos automatically display information for students and reduce students' motivation to actively seek to understand what is not clear to them. This relationship suggests that the adaptive learning environment may play an important role in providing additional explanation, but it also reduces students' extraneous load and makes managing intrinsic load easier, possibly at the cost of deliberate difficulty.

3.5.2 Frustration Linked to Teacher Energy and Lack of Student Knowledge

Findings seem to indicate that students' frustration with the adaptive learning experience may be linked to the monotony of the video instructor or the students' own lack of content knowledge. In regard to the theme of frustration, among the 12 students (n = 12) who did not achieve a perfect score on the final test, 50% of those students found the adaptive learning session frustrating.

When asked in the interview, half of the participants said that they felt frustrated. I found two different reasons why students were frustrated: (1) they did not like the level of energy (i.e., monotone and low enthusiasm) of the teacher in the video and (2) they lacked knowledge about the adaptive tasks they were tested on. Some of the comments about frustration given by students were:

> *"I was frustrated when he (the video instructor) took a while to explain things."*
>
> *"This video (...) was like boring, it was like slow sort of, but not really due to the content."*
>
> *"When the quizzes were talking about something else (I had not learned yet), it was kind of confusing at first."*

Similarly, among the six students who performed perfectly on the final test, 50% (n = 3) found the adaptive learning session frustrating because it tested them on knowledge that they did not yet possess. Both groups (frustrated and not frustrated) stated that, after they watched the remediation videos that explained the missing concept thoroughly, they felt more confident about their learning process.

The findings suggest that students got frustrated engaging in the adaptive learning activity when they did not know the answers to questions about a concept that they had not learned about. In particular, the questions on the quiz preceding the adaptive task were designed in such a way that students who had strong backgrounds in mathematics and/or economics would be able to successfully perform the adaptive tasks (the quiz) but those without this background would not.. ; In a real setting, the high-scoring students would not have been given the adaptive learning task, but in this simulated experiment, all students received it regardless of their scores on the pre-remediation quiz. The adaptive learning procedure may have created student frustration because the initial instructions were not clear; they may not have been aware that the pre-remediation test was intended to test them on knowledge that they were to receive but were not yet expected to possess.

Since in this study the enjoyment of video and learning gains are not positively correlated, it seems worthwhile to examine whether other factors included here might be correlated. According to one study, researchers found that engaged concentration and frustration are correlated with positive learning outcomes (Pardos, Baker, San Pedro, Gowda, & Gowda, 2013); therefore, I might examine ways to balance students' enjoyment of a learning process and the increased frustration necessary to evoke a student's positive learning gains. Furthermore, it is valuable to explore what factors in adaptive learning environments create frustrations that add value to students' learning gains, as well as what factors do not add value. According to Guo et al. (2014), using a conversational, enthusiastic teaching style enhances students' engagement. The current study results seem to support their conclusion in that some students found the adaptive learning session frustrating because the energy of the teacher was low or that the style was not appealing to them. Thus, to reduce students' frustration, videos with a conversational and enthusiastic style should be selected for use in adaptive learning environments.

3.5.3 High Rates of Online Video Usage as Educational Supplements

The interview data seem to suggest that students use online videos to supplement their understanding of topics they are learning about in the classroom. From the interview data, in regard to online education, 94.4% of the participants (n = 17) indicated that they use MOOCs or YouTube Education outside of school courses to help them understand concepts.

Some students explained how they used online educational materials:

> *"I watch Khan and YouTube videos. I take like bits and pieces (of) knowledge that I need help (with) for school."*

"If I am looking for a specific topic that I don't understand, I just search on YouTube instead of having to browse through an entire (set of) notes."

"If a professor doesn't really explain it (a specific topic) all the way, I prefer using online videos because I can pause (the videos and watch) it over (again)."

Students seem to regularly use online videos as supplemental learning tools. In a descriptive and experimental study on college students (n = 91), Jaffar (2012) demonstrates that 98% of respondents were using online learning videos as a source of information. Even though my study is based on a smaller sample size (n = 18), my result supports Jaffar's result by showing that more than 90% of the study subjects are using MOOCs or YouTube Education outside of school. Results from this study may reinforce Jaffar's claim that online video education has become essential for new learners in their undergraduate learning. Therefore, teachers and online course designers should also increase their efforts to continuously improve online teaching quality in order to help the current and upcoming generations of students.

3.5.4 Limitations

This study also has some limitations. First, the sample size was very small. However, as a pilot study, this study can assist other researchers in determining an experimental research design for future research on frustration, enjoyment, and the use of online educational videos. Second, this study is limited in its analysis of the behavior of learners who participated in an adaptive learning activity featuring MOOC material from a single lecture. Viewing or posting comments, which are often part of learners' full MOOCs experience, were not considered. Future research should also examine how MOOCs' forums and posted comments may also play roles in affecting students' adaptive learning experiences. Third, this study focused on using thematic analysis of the qualitative data. Future research studies could consider employing machine learning techniques alongside thematic analysis to analyze the frustration and engagement of learners in adaptive learning environments.

3.6 Conclusions and Future Work

Prior research suggested that behavioral patterns such as boredom or frustration in adaptive tasks should be explored in order to improve students' learning experiences (Rosen et al., 2018). Using collected survey data, interview data, and post-test scores from the 18 students in my experiment, this study investigated engineering undergraduate students' perceptions of an adaptive learning environment. After the data collection, an analysis of the descriptive statistics and interview data were used to identify emergent themes. In this experiment, the results suggest that there may be a negative correlation between students' learning gains and their perceptions of enjoying an adaptive task.

Several insights gained from this pilot study may help to inform the design of future research studies of adaptive learning experiences. First, I learned that for future research studies, I should administer a pre-test to all students before they engage in the adaptive lecture lesson. This will allow us to compare learning gains across students who perform various adaptive tasks. Also, future research designs should use a machine learning model to better direct the learners' path through instructional materials. In addition, detailed qualitative data collected through open-ended interview questions should be used to better understand the quantitative data and the survey results.

Chapter 4 Methods

4.1 Description of Experiment

The primary purpose of this exploratory research study is to identify factors that affect undergraduate engineering student learning and experience in an online learning environment with interventions. To achieve this purpose, I took several steps. First, I created online learning modules in collaboration with a professional lecturer and an experimental procedure. The experimental procedure consisted of having participants receive the prepared online modules, take content quizzes, and complete survey questions, and conducting semi-structured interviews. Second, I recruited undergraduate engineering student participants (N=70) who were unfamiliar with the intended procedure topic and I administered the experimental procedure to them. After collecting the data, I used several methods to analyze the data. Specifically, I used descriptive statistics about the study population, and I analyzed the quantitative data using multiple linear regression analysis. I then used thematic analysis, an approach common to a great deal of educational and psychological research, to analyze the qualitative data. Finally, I merged the quantitative and qualitative findings to mutually illuminate the two types of data in order to understand how each type of data informs the other. Figure 4-a below displays the overall design of the experiment.

Figure 4-a: The Overall Design of The Experiment

4.2 Location and Equipment

The experiment took place in the Engineering Research Building Laboratory at the University of Michigan. The participant station consisted of a MacBook Pro "Core i5" 2.7 13-inch laptop running in macOS Mojave.

4.3 Materials

For the experimental design, I created two sets of materials: (1) the modules (the online teaching materials) and (2) the online learning environment (a series of procedures that students complete). I created these materials at a level of difficulty that I thought was appropriate for undergraduate engineering students in that the materials were hard enough but not too hard, i.e., challenging to some degree, but accessible to students with an understanding of basic mathematics and statistics, in my estimation. It should be noted that the possible effects having to do with pure language concerns, e.g., effects of English as a second language, on the part of either the lecturers in the materials or the participants, have not been specifically studied in this experiment. The potential effects have been treated as negligible for this study; however, some discussion touches upon this concern later on. For this reason, I reported citizenship status in the Descriptive Statistics section because of the higher likelihood that participants with the second language concerns might be non-citizens, but I have not explored that in depth in this research. I intend to incorporate this consideration in subsequent studies.

4.3.1 Module Materials

Materials used in these experiments were developed according to current best strategies and current applications of cognitive load theory. Studies examining the process of developing learning materials suggest that materials be segmented so as not to overload students with information (Guo et al., 2014). For example, the study shows that too long videos will create extraneousness (cognitive load that does not have a payoff in learning) and that 3-6-minute videos are optimal length. Based on this information about the video, other forms of materials were created to give the same information to the students.

Another strategy for development is matching modality, which is a strategy to use materials appropriately to activate visual and auditory channels optimally to produce a better experience (Mayer & Moreno, 2003). I considered this when I chose audio, visual, and text materials. To the extent possible, I also employed the strategy of weeding, which ensures that students can focus on the material without background noise and distracting visuals that might confuse them (Ibrahim et al., 2012). I could not alter the video materials and visual aids, but I tried my best to weed out any unnecessary audio or visuals.

For the text materials, I also used the strategy of signaling (De Koning, Tabbers, Rikers, & Paas, 2009), which is the use of highlighting, underlining or bolding to let the learner know what is important.

In this subsection, I explain what materials I created and my purposes for creating them in the form that I did. They are (1) primary teaching videos; (2) interventions; (3) test and quiz questions; (4) experience surveys; and (5) a demographic survey.

(1) Primary Teaching Video: To create the online modules, I used clips of online recorded videos taken from a MOOC class on Operations Research. The instructor, Dr. David Mendez, an associate professor in the Department of Health Management and Policy at the University of Michigan School of Public Health, provided his class materials from HMP 654: *Operations Research and Control Systems* on the topic of decision analysis. Dr. Mendez and I chose decision analysis because (1) it is an engineering topic that any student with basic probability knowledge can understand; (2) freshman and sophomore undergraduate students (who are my main target subjects) typically have not been exposed to the topic; and (3) even junior and senior-level industrial and operations engineering students usually have not been much exposed to the topic.

    In the video decision analysis, participants learned about three types of individual subjects: (1) the risk-neutral decision-maker, who is indifferent to risk when making an investment decision; (2) the risk-seeking decision-maker, who prefers an investment with an uncertain outcome over one with the same expected returns and certainty that they will be delivered; and (3) the risk-averse decision-maker, who prefers certainty and does not like betting on uncertain outcomes. The primary teaching video was segmented into three 3-6-minute videos, one for each subtopic.

(2) Interventions: Five types of interventions [(a) four interventions that utilize the two learning channels to varying extents and (b) a writing reflection intervention] were generated. These choices were derived from the findings of the pilot study discussed in Chapter 3 and the literature review in Chapter 2.

   a. Four types of interventions that utilize the two learning channels: I created four different types of interventions that utilize the two learning channels to varying extents, and Dr. Mendez double-checked, revised, and did final approval of the intervention materials.

   - The Audio-only intervention material took the form of 3-6 minutes of audio lectures. Thus, this intervention uses primarily the auditory channel. However, a figure is included because of the complex nature of the information presented.
   - The Text-only intervention material consisted of approximately a half-page of text accompanied by a figure (again necessary because of the complex nature of the information presented). Thus, this intervention uses primarily the visual channel.
   - The Video intervention material also took the form of 3-6-minutes of video lectures but in a style different from that of the main teaching video. Thus, this intervention uses both the auditory and visual channels.
   - The Video+Text material consisted of three different combinations of the video and text interventions just described. Thus, this intervention also uses both the auditory and visual channels.

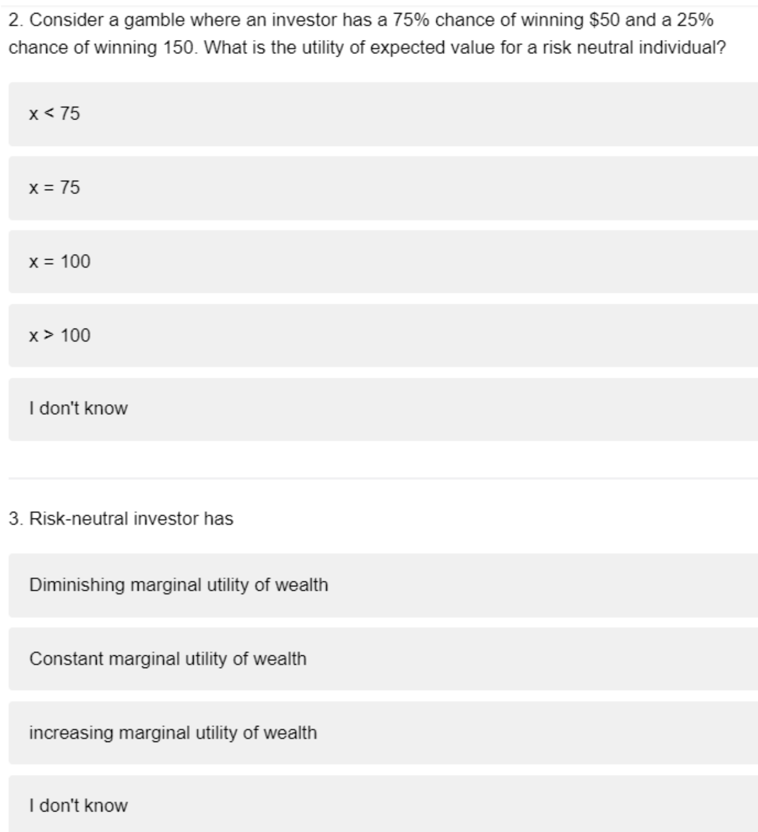
b. Writing reflection intervention: I created a writing reflection intervention consisting of a screen asking students to write one to two sentences summarizing the topic that they had learned about.

Sample screenshots of all of these interventions are available in Appendix A.

(3) Test and Quiz Questions: I created pre-test and post-test questions and content quiz questions based on the course material provided, and Dr. Mendez double-checked, revised, and did final approval of the questions so that the difficulty level was appropriate for undergraduate engineering students new to the topic.

a. Pre-test & Post-test: The purpose of these tests was to measure their existing (before and after) knowledge about the topic. Both tests consisted of questions about the topics presented: definitions of terms. probability questions, and so on. Pre-test scores were used in conjunction with post-test scores to measure students' learning gains about the topic.

b. Content Quiz: The content quizzes consisted of four questions. These content quizzes were similar in style to what the student saw in the pre-test and post-test, but the questions were different. Sample questions are shown in Figure 4-b below.



2. Consider a gamble where an investor has a 75% chance of winning $50 and a 25% chance of winning 150. What is the utility of expected value for a risk neutral individual?

- x < 75
- x = 75
- x = 100
- x > 100
- I don't know

3. Risk-neutral investor has

- Diminishing marginal utility of wealth
- Constant marginal utility of wealth
- increasing marginal utility of wealth
- I don't know

Figure 4-b: Sample Content Quiz Questions

(4) Experience Survey: I created six 5-point Likert scale survey questions about students' perceptions (self-reported memory, self-reported knowledge gain, and desirable difficulty). The six experience survey questions developed in this study were modified from the work conducted by Pomales-Garcia and Liu (2006) to collect students' perceived experiences as quantifiable data. Two of these questions were:

- Before the modules, how much did you know about the topic discussed in the module, using a scale of 1-5 where 1 = completely new material and 5 = expert?
- After the modules, how much did you know about the topic discussed in the module, using a scale of 1-5 where 1 = completely new material and 5 = expert?
- How frustrated are you with this online learning task (quiz) using a scale where 1 = Not frustrated at all, 5 = Very frustrated? (please note that there were no labels attached

for 2, 3, and 4) (footnote: For this study, I was attempting to make it easier for students to rate their perception, and I relied on students' familiarity with Likert scales . However, for future study, I would like to label each number, which will enable me to quantify the qualitative experience more precisely)

(5) Demographic Survey: I created a demographic survey that collected information about the following: students' sex, age, major, GPA, citizenship status, and race.

4.3.2 Online Learning Environment

I created an online learning environment using Qualtrics software. The environment consisted of a combination of the materials mentioned above. Students entered the environment with Topic 1. No supporting materials were given or permitted. Participants spent approximately 20 to 30 minutes participating in the quantitative portion of the experiment.

(1) Pre-test: Participants were given a pre-test about Topic 1. There was no time constraint on the pre-test.

(2) Main Teaching Video: Participants were given a 3-6-minute video on Topic 1. They did not have the capability of jumping ahead or returning to previous sections.

(3) Content Quiz: After the video, participants next encountered a content quiz about Topic 1. There was no time constraint on taking the quiz.

(4) Interventions: Participants next received one of the four interventions described in the Materials Section 4.3.1. In a real-world online learning environment, only students with scores below a threshold preset by the lecturer or determined by machine learning would receive an intervention after the quiz. However, this experiment was designed to simulate cases in online learning environments where the students receive the intervention;

therefore, each participant was provided with an intervention after each subtopic presentation.

(5) After the intervention, half of the participants (n = 35) received an additional intervention, a writing reflection task, in which they had to write about what they had just heard, watched, or read. The distribution of participants receiving the writing reflection intervention was predetermined by the order in which the participants joined the trials.

(6) Post-test: Following each intervention, all participants were given a post-test about Topic 1. There was no time constraint on the post-test.

(7) Experience Survey: After the post-test, participants received six 5-point Likert scale survey questions in Qualtrics about their perceptions (self-reported memory, self-reported knowledge gain, and desirable difficulty).

After finishing the Topic 1 segment, the students were directed through the same procedures for Topics 2 and 3. However, each time students went through the cycle, they received a different type of intervention. For example, the first quarter of students received first an audio intervention, second a text intervention, and third a video intervention. below displays the intervention path that each group of students followed.

Table 4-1: Intervention Paths

	Intervention Tested		
Participants	Trial 1	Trial 2	Trial 3
1-18	A	T	V
19-36	T	V	V&T
37-53	V	V&T	A
54-70	V&T	A	T

A: Audio, T: Text, V: Video, V&T: Video & Text

(8) After the third lecture procedure, the system displayed a set of survey items to collect demographic information about the participants, as noted above.

4.4 Recruitment and Selection of Participants

For this study, I recruited 70 participants. Initially, I had planned to recruit only 60 students, which was the number required to collect 30 data sets for each factor examined. However, I subsequently recruited ten additional participants to provide a richer dataset. Each study participant was 18 years old or older and currently an undergraduate engineering student at the University of Michigan (UM). Undergraduate students were selected because, as noted above, they could be assumed to have minimal knowledge about the lecture topic, specifically decision analysis. The UM Institutional Review Board (IRB) reviewed this study and determined that the research was exempt (HUM00161103); participants would need only to be orally informed about the experimental procedure, the benefits of the study, the potential risks, the compensation, the confidentiality of the study record, and the voluntary nature of the experiment. After receiving the exemption, I contacted each UM engineering department's administrator, who then sent my recruitment email, which contained screening criteria (see Appendix B), to the students in his or her department. Instead of sending emails to all engineering departments at once, I recruited the participants using rolling enrollment. As compensation, participants received a $20 MasterCard gift card for participation in this study. The recruitment period was from April 10, 2019, to July 30, 2019.

4.4.1 Descriptive Statistics—Characteristics of the Sample Population

My email received a total of 70 responses. The ages of the respondents ranged from 18 to 32, with a mean age of 20.69 (SD = 1.96). Of these respondents, 58.6% (n = 41) were males and 41.4% (n = 29) were females. The average GPA was 3.42/4.00. The racial makeup of the sample was 62.86% White/Non-Hispanic, 24.29% Asian, 5.71% Black or African American, and 7.14% Hispanic or Latinx. This racial makeup is similar to the racial makeup of the University of Michigan's engineering student enrollment (Michigan, 2020). The study population comprised 90.00% United States citizens, 7.14% neither United States citizens nor permanent residents, and 2.86% United States permanent residents. As shown in

Table A-2 in Appendix C, the majority of participants were from the mechanical engineering and industrial engineering departments.

4.5 Data Collection

For this work, the bulk of the data are quantitative, and these were obtained through the pre- and post-tests, content quizzes, and surveys. However, a significant portion is qualitative and was obtained through the individual interviews. I scheduled a one-hour session with each student, and the data were collected over a period of four months from April 10, 2019, to July 30, 2019.

The overall data collection procedure for this research study consisted of three steps: (1) after being informed about the study procedure and giving verbal consent, participants engaged in the online learning environment and received interventions; (2) participants' demographic information was collected; and (3) participants were interviewed about their online learning experience. The steps of the quantitative data collection are illustrated in Table 4-2, which represents the repeated procedure for data collection for the three lecture topics in the modules plus the demographic survey at the end. These test scores were recorded manually so that later I would be able to use the data in the analysis.

Table 4-2: Procedure for Quantitative Data Collection

Step	Description	Output Collected
1	Pre-test	Pre-test Scores
2	Main Video Lecture	None
3	Content Quiz	Content Quiz Scores
4	Interventions	What type of Intervention
5	Writing Reflection	Did or did not do
6	Post-test	Post-test Scores
7	Experience Survey	Likert Scale Scores
8	Demographic Survey	Major, Gender, Race, etc.

$\left.\right\}$ For each topic

After the demographic survey, in order to supplement the quantitative data and provide crucial context about participants' experiences in the experiment, I conducted interviews. I designed the interview questions as a modified version of the questions in my 2018 pilot study (Kwak & Mondisa, in review; these are listed in Appendix D). The interview questions asked about three themes: frustration, attention level, and enjoyment of the material. These themes were chosen because Rosen et al. (2018) suggested that behavioral patterns such as boredom or frustration in online tasks should be explored in order to inform the design of better online learning environments. By conducting interviews, I was able to gather complex, in-depth data about students' perspectives that would not have been easily obtained through questionnaires or yes-no interview approaches. As the work of Creswell (2002) shows, interviews enable researchers to probe for further information, elaboration, and clarification of responses; and for the participants, interviews offered a "feeling of openness" to their responses.

4.6 Data Analysis

I used a mixed-methods research approach to identify factors that affect students' learning in an online learning environment. Mixed methods research is the use of both quantitative and qualitative data analyses in a research study to investigate the data via statistical and mathematical analysis and analytical examination of data obtained directly from participants by way of surveys, interviews, etc. (Creswell, 2002). In my study, I used qualitative and quantitative results to arrive at the findings. The analysis of the data did not start until all students had completed the experiment.

4.6.1 Data Analysis Procedures—Quantitative

I used multiple linear regression analysis to determine the degree to which independent variables (e.g., demographic information, intervention type, delivery of information, pedagogical approach) were able to predict the students' learning gains. I used Tukey's tests to find statistical differences between the means of the learning gain scores among interventions; I also used it for the mean of the level of frustration with intervention. A software package, IBM SPSS 25, was used to perform both analyses.

4.6.1.1 Multiple Linear Regression Analysis

Multiple linear regression analysis was used to investigate the relationship between the individual independent variables and the variance in the dependent variables (learning gains and self-reported memory). Multiple linear regression is a useful tool because it can statistically describe the significance (coefficients) of independent variables to changes in the dependent variable (Meyers, Gamst, & Guarino, 2016). In this study, it was used to find out which independent variables best predicted learning gains and find their relative contribution to the learning gain.

In the multiple linear regression analyses, the level of significance was set as $\alpha=0.05$. The stepwise method was used to identify the degree to which independent variables predicted learning gains[2]. I was interested in several components from the outputs of the multiple regression, specifically the R^2, F-value, beta coefficient, and p-value. The R^2 is used to statistically interpret the degree of relationship between combinations of the independent variables and the dependent variables (Meyers et al., 2016). The F-values and p-values are used to determine whether the relationships between the sets of independent variables are significant (Meyers et al., 2016). The beta coefficient (β) is used to find the contributions from each of the independent variables to changes in the dependent variables (Newton & Rudestam, 1999).

4.6.1.2 Tukey's Test

I used Tukey's tests for two purposes: (1) to find statistical differences between mean learning gain scores across four interventions, and (2) to find statistical differences between levels of frustration with interventions across the interventions. Tukey's test statistics compare each mean with the other three means. In the test, the level of significance was set as $\alpha=0.05$.

4.6.2 Data Analysis Procedures—Qualitative

In my qualitative analysis, I used thematic analysis to uncover trends in thoughts and opinions that could provide greater insight into the quantitative outcomes. In order to check the legitimacy and increase the reliability of the thematic analysis, I used interrater reliability tests, which I conducted with a research assistant.

4.6.2.1 Thematic Analysis

Thematic analysis is a method for examining interviewees' words in the transcript to identify key patterns of emerging elements within the data and then analyzing those elements

[2] I attempted to use multiple linear regression analysis for self-reported memory, but the result did not include any usable data (see the discussion section).

(repeated keywords and phrases) to form them into codes (Braun & Clarke, 2006). Codes, in qualitative analysis, are defined as "tags or labels for assigning units of meaning to the descriptive or inferential[3] information compiled during a study" (Miles, Huberman, & Saldaña, 1994). The set of codes makes up a codebook that the investigator compiles to enable him/her to analyze the corpus of transcript data.

As noted above, the process of thematic analysis enabled me to extract complex relationships among students' descriptions and ideas about their experiment experience. I conducted line-by-line coding, in which I looked at every line and labeled everything I believed to be significant. For the coding process, I used a software package, QSR NVivo 10, that allows the user to easily label quotations in the interview transcripts and attach them to a code or codes. The term *quotation* can refer to a word, a phrase, a sentence, or a group of sentences. To generate the codebook for the whole qualitative analysis, I followed a series of six steps. First, I used a professional transcribing service to produce the transcripts. Second, I found the sample from which to generate the preliminary codebook. Third, I generated the initial codebook from the sample that I had selected. Fourth, I reviewed and revised the code in the context of the data, following thematic analysis steps developed by Braun and Clarke (2006). Fifth, on the basis of the initial data, I conducted interrater reliability tests on the preliminary codes with another researcher to validate the legitimacy of the initial codebook. Lastly, I coded the remainder of the transcripts using the validated codebook. The thematic analysis enabled me to focus on what was significant and on common patterns that were occurring.

[3] By inferential, I mean, for example, I am attaching the tag "difficult intervention" to expressions that I can reasonably interpret as related to difficulty, e.g. ("intervention was kind of hard to follow" or "that intervention was easier to get distracted.")

4.6.2.1.1 Audio Transcribing

Audio interview data were transcribed by a professional transcribing service called Rev. I strictly confined the basis of the code to the words as transcribed by the professional transcribers. The transcripts that were received did not contain suprasegmental information (pauses, emphasis on certain words, intonation, etc.).

4.6.2.1.2 Choosing the Sample

In choosing the sample to make the preliminary codebook, I wanted to extract the codes from samples of the students' transcripts that reflected all the possible online experiences. To do so, I looked at how to categorize the transcripts so that all the experiences were represented. To get a representative sample, I divided the transcripts into 8 subgroups by what path of interventions they received and whether or not they had the writing reflection intervention, as shown in Table 4-3 below. I wanted to make sure that I got a transcript from each of the subgroups that I would get by dividing the whole group up according to these categories.

Table 4-3: Eight Subgroups by Types of Interventions Received

Writing Reflection	**No Writing Reflection**
A-V-T	A-V-T
V-T-VT	V-T-VT
T-VT-A	T-VT-A
VT-A-V	VT-A-V

I then took two random samples from each subgroup because I estimated that only one would probably not adequately represent each subgroup. Therefore, I generated the preliminary codebook from transcripts provided by 16 students.

4.6.2.1.3 Generating a Preliminary Codebook from Samples

As preparation for analyzing the whole corpus, I generated a preliminary codebook from the transcripts of the initial selection of 16 transcripts. On the basis of Braun and Clarke (2006), I first created a 6-step procedure for analyzing the transcripts. I then used thematic analysis to analyze the initial transcripts, as shown in Table 4-4 below.

Table 4-4: Six-Step Procedure for Analyzing the Transcripts

Thematic Analysis Steps (Braun & Clarke, 2006)	**How I Performed the Data Analysis Step**	**Result/Output**
1. Familiarizing myself with the data	Read and re-read a second time the sample transcribed data. Made notes about words or phrases occurring in a particular pattern	Log of preliminary ideas
2. Generating initial codes	Coded recurrent or unusual features of the data in a systematic fashion across the entire sample data set	Log of data relevant to each preliminary code
3. Searching for themes	Organized codes into potential themes	Log of data relevant to each potential theme
4. Reviewing themes	Checked to see whether the themes worked in relation to the coded extracts	Checklist of reviewed themes
5. Defining and naming themes	Iterated the analysis to refine the specifics of each theme	List of clear definitions and names for each theme
6. Producing the report[4]	Selected examples relating the analysis to the research question and literature (the final analysis of selected extracts)	Report on the analysis and the codebook

In the following paragraphs, I provide a description of the steps of my application of thematic analysis to my corpus:

[4] The 6th step is reserved for the analysis including the remainder of the transcripts.

Step 1: Familiarizing myself with the data. I read the selected 16 transcripts (as noted above) twice and wrote down the initial emerging ideas. Table 4-5 below displays an excerpt from the log of initial ideas I created:

Table 4-5: Initial Ideas Sample

Initial Ideas
• Engineering students are visual learners. • Video plus text is the best intervention. • The text could be providing unnecessary information. • Students seem to regard audio interventions as the worst interventions. • Audio is confusing to students; students don't know what the audio is talking about. • Text interventions allow students to progress at their own pace. • Video is good but video plus text is better.

Step 2: Generating initial codes. I systematically tagged recurrent words and phrases throughout the body of sample transcripts and collected examples related to each potential code. Table 4-6 below displays an excerpt of the log of codes I created. Note that items in initial codes are highly abbreviated and are much more fully articulated in later steps.

Step 3: Searching for themes. After collecting all the information relevant to each code, I organized these initial codes into prospective themes. For example, I assigned codes such as "Audio was confusing," and "Audio was harder" to the preliminary theme "Audio was worst". All these themes were compiled into a log.

Step 4: Reviewing themes. I checked to see whether the themes worked in relation to the coded extracts in the sample data set (and, subsequently, throughout the entire data set). This process provided me with potential insights about overarching themes encompassing the identified preliminary themes and codes (e.g., "desirable difficulty" and "non-value-adding frustration"). I generated a checklist of reviewed themes.

Table 4-6: Excerpt of the Log of Codes Created

Initial Idea	Parti-cipant #	Sample responses to an interview question: 'How would you rank interventions you were given that affected your learning? And why?'	Initial Codes
• Engineering students are visual learners. • Video plus text is the best. • Text could be just unnecessary information. • Audio is the worst. • Audio is confusing. Students don't know what the audio is talking about. • Text could be helpful to students going at their own pace. • Video is good but video plus text is better.	1	The audio was least helpful. Video was probably second. Video and the text first. Because like personally I'm more of a, a visual learner. So like audio, it was really hard to know what he was talking about. With the video was really helpful because they would draw something on the picture or you could even see the mouse reference what they're, what they're indicating.	• Video + text best • Audio least helpful • Video second • Video first • Visual learners • Drawing & moving the cursor were helpful for references
	2	I would say video and the text was best. I feel like they were kind of similar. The video and then the video and the text were similar, because I just watched the video for both of them, and then I went and looked at the text after. I was like, "Oh, that's what I just learned." I felt like the text was repetitive, so I just skipped. The audio was obviously the worst. it was a little harder, because I feel like ... if I remember correctly he was saying like, "And here this thing moves," and I'm like, "I don't know what you're talking about."	• (Video and Video + Text) Similar • Repetitive • Skipped • Audio worst • Harder • Don't know what it is talking about
	3	I liked the video and the text is number one. Videos second and then the last was audio. It's just a good supplementary ... I think what would it look like for the texts, like it was like a transcribed version of the video. But I think if instead of it being the transcribers on the videos, like more supplementary, like a couple different examples, it would be much better. But that sort of thing helped me. I had more time, I wasn't pressured by the time limit in the video. I could just sit there and read at my own pace and think about what I saw in the video and then reflect that back on. The way I operate is more visual and then influencing. I like see something and I like see it a few more times before it actually makes sense and I ask some small questions here and there, and I'm that's just worked for me so far. And so with the audio, I saw the picture and I knew what he was talking about, but sometimes he'd be like, "Oh the expected value of this, a little bit lower than that." And I'm like, "I don't know what you're talking about."	•Supplementary • Different example • Pressured • Time limit • Own pace • Visual learner • Confusing

Step 5: Defining and naming themes. I continued to refine the specifics of each theme and the general analysis. I produced a clear definition and name for each theme.

Step 6: Producing the Report Codebook[5]. I selected vivid, compelling extract examples, performed the final analysis of selected extracts, related the analysis back to the research questions and literature, and produced a scholarly report of the analysis (see Section 5).

4.6.2.1.4 Interrater Reliability Testing Procedure

I performed an interrater reliability testing procedure to validate my qualitative data analysis. Interrater reliability, in qualitative analysis, is a calculation of the internal consistency of a group of different qualitative raters (Gravetter & Forzano, 2018). Research suggests that a coding done by a single researcher may be subject to unrecognized mistakes and/or bias that go on throughout the analysis; however, if another person acts as a rater to produce the same analysis, consistency between the two raters will tell us that we can be more certain that any other rater(s) who interpret the data will come to the same conclusions (Marques & McCall, 2005).

In this study, interrater reliability testing was used to compare the decisions made by the researcher (the author of this study) with decisions made by a different rater (the research assistant) to judge the validity[6] of the coder's decisions. I took the following interrater reliability test steps:

1. I read and coded two transcripts from each of eight participant groups (n = 16 interview transcripts) to develop preliminary codes and a codebook.

[5] As previously explained, in the codebook generation procedure, this step consisted of producing a codebook rather than a report

[6] Validity is tested with respect to usability, accuracy, appropriateness, and agreement with the coder's decisions.

2. I trained an undergraduate researcher to prepare her to serve as a rater, as explained in greater detail in the section below. Training consisted of (1) giving instructions for using NVivo software and explaining the coding process, and (2) the interrater reliability test procedure
3. I provided the rater with the initial codebook and had her independently apply the code to the same four transcripts.
4. The rater and I met to compare our results.
5. We clarified our disagreements about the codes identified; we also captured additional codes that we identified and added them to the codebook. This process was conducted until 100% agreement on the codes was reached.
6. Both the rater and I evaluated four more transcripts using the updated codebook. Again, we compared our coding and discussed our disagreements about the codes and descriptions until 100% agreement was reached. This process was repeated for four transcripts at a time, until the 16 sample transcripts were coded.
7. Finally, together, the rater and I went over the codebook and clarified our codes/themes and the corresponding definitions and answered any questions the rater raised so that I could establish the final codebook.
8. I used the final codebook to code the remaining 54 interview transcripts.

4.6.3 Interrater Reliability Sessions

As indicated above, we achieved 100% agreement through a series of discussions and several interrater reliability iterations. In the following sections, I describe these in greater detail.

4.6.3.1.1 Interrater Reliability Rater Training Session

After the creation of the preliminary codebook that I generated with 16 interviews, I trained one undergraduate research assistant (RA) on how to use the codebook to code interviews. Initially, I provided my research assistant with the codebook and four interview transcripts. I also provided her QSR NVivo 10 software to conduct qualitative research and a tutorial on how to use the software. I provided her with the following instructions:

1. Read a transcript and assign code(s) to sentences or phrases that reflect themes in the codebook based on her interpretation.
2. Critically question my interpretations of the codes and be prepared for me to question her interpretations.
3. In situations when two or more codes apply to a specific segment of text, assign multiple codes to them.

4.6.4 Refining the Matching of Code and Quotations

Thorough comparison and discussion of our results are important in achieving reliability of the codebook. After the RA and I each evaluated the coding of initial interviews 1-4, we identified which quotations should be matched to each code. After the completion of the coding, we had a set of quotations that were highlighted and associated tentatively with particular codes. For each of these quotations, we considered two possibilities. First, if we agreed that the quotation matched the code, we moved on to the next available identified quotation. If we did not agree that the quotation matched the code, we each explained our reasoning. We discussed each mismatched code and quotation until we reached the same conclusion for all of them. If we agreed that a quotation did not match the code, then we removed the quotation from the codebook. We iterated this process until we reached 100% agreement on all code and quotation associations.

4.6.5 Refining the Code and Definition

After the code and quotation matching refinement process, we refined the definitions of some of the codes to be more specific. For example, we originally used the code "Conversational Style" for quotations that describe how the conversational style of the teacher influences the learning of the individual student. The RA and I discussed whether in this context the 'Conversational Style' code referred only to the online lecturers whom participants watched, rather than to participants' general idea of conversational style of lecturers. We refined the meaning of 'Conversational Style' to refer to what an individual student said about how the conversational style of the teacher in the online setting influenced his or her learning. Also, we renamed the code 'Online Conversational Style' to distinguish participants' observations of online conversational styles from those in other settings. We proceeded to discuss and refine all other disputed codes in the same way. We then updated the codebook with the new definitions.

With the updated codebook, the RA and I conducted three more rounds of interrater evaluation for Interviews 5-8, 9-12, and 13-16. For each set of interviews, we conducted the code and quotation matching refinement and code and definition refinement processes. We performed these processes until we reached 100% agreement. These four iterations of the interrater reliability process were completed to ensure an accurate codebook and definitions.

Once this interrater reliability test had ensured that the codebook and the analyses appeared to be valid, I used the finalized codebook to code 54 interviews. After coding all 70 interviews, I was able to develop well-defined themes. The steps are somewhat fixed and generic but what actually happened in the sessions is particular to this case.

4.6.6 Merging Qualitative and Quantitative Data & Interpretation

At this point, I performed combined qualitative and quantitative analyses to achieve a better understanding of the factors affecting undergraduate engineering learning and experience in the online learning environment.

For Research Question #1, I wanted to explore what factors affect students' learning gains in an online learning environment. Quantitatively, I used a stepwise multi-regression analysis to find which variables were significant in leading to increased learning gains for students. Qualitatively, I interviewed the students and asked them to explain what elements helped them with their online learning. I used thematic analysis to summarize themes from the interview data. Finally, I compared and contrasted my quantitative data variables to the qualitative data to summarize my results and explore the answers in depth.

For Research Question #2, I again followed the aforementioned process. This entailed interpreting and summarizing the effects of factors in the environments through quantitative (i.e., Tukey's test and descriptive statistics) and qualitative (i.e., interview data coding about self-reported memory) results. I then refined and identified the emergent themes and investigated my hypotheses.

For Research Question #3, I again followed the aforementioned process. This entailed interpreting and summarizing the effects of online learning environments through quantitative (i.e., multiple linear regression analysis) and qualitative (i.e., interview data coding about self-reported memory) results. I then refined and identified the emergent themes and investigated my hypotheses. In the following chapter, I report results and findings from this experiment.

Chapter 5 Results

The purpose of this study is to examine what factors (e.g., types of intervention, presentation of content, activities) affect undergraduate engineering students' learning gains, learning experience, and self-reported memory in an online learning environment. The study was conducted to identify data that will aid online education designers to better design an online learning environment, which is important to ensure high quality. This section reports the study's quantitative and qualitative findings with respect to each research question. Table 5-1 and Table 5-2 display summaries of the quantitative and qualitative results, respectively, found for each research question. I first present the quantitative results. After that, I present the qualitative results.

Table 5-1: Summary of Quantitative Results

Summary of Quantitative Data Results		
Research Question	**Trial**	**Significant Factors**
RQ1: What factors in online learning environments affect *learning gains* (i.e., measured difference between the post-test and the pre-test scores) for undergraduate engineering students?	1	Pre-Test Score ($p < 0.01$) Performance of Writing Intervention ($p < 0.01$) Level of Frustration with Intervention ($p < 0.01$)
	2	Pre-Test Score ($p < 0.01$)
	3	Pre-Test Score ($p < 0.01$)

Table 5-2: Summary of Qualitative Results

Research Questions	Interventions Types/Factors from General Findings	Major Theme	Explanation of Theme
		Summary of Qualitative Data Results	
RQ2: What factors in online learning environments affect the *learning experience* for undergraduate engineering students, and, specifically, what factors produce desirable difficulty?	Audio-only Intervention	**1.** Audio Intervention Not Visual, Not Helpful	Audio intervention presented to them was not visual enough and thus it was not helpful to them.
	Text-only Intervention	**2.** Information Quantity Overwhelming, Yet Allowed Self-pacing	The quantity of information in the text intervention presented to them was overwhelming but students could process the information at their own pace.
	Video Intervention	**3.** Seeing the Lecturer's Engagement with Materials Helpful	The visual aspect of the video intervention enabled students to see the lecturer engage with the material through actions.
	Video+Text Intervention	**4.** Variety in Choice of Interventions Helpful, Provides Supplements	The mixed intervention enabled students to choose among various types of learning interventions, and the other types of intervention that they did not choose initially acted as a supplementary intervention.
	Writing Reflection Intervention	**5.** Writing in Own Words Useful and Engaging	Writing in their own words (paraphrasing and summarizing) supports learning through engagement.
	Inconsistency in Terminology	**6.** Efforts to Understand Terminology Helpful for Understanding	Putting in the effort to figure out the connections or relationships between discrepant terms helped some students understand the material.
	Repetition	**7.** Repetitious Format Both Annoying and Helpful	Repetition in the online learning environment created annoyance; some were also judged to be helpful by others.
RQ3: What are the factors in online learning that affect undergraduate engineering students' *self-reported memory*?	Interactive Tasks	**8.** Interactive Task Participation Supports Learning, Confidence	Participation in interactive tasks positively influenced their learning and made them more confident about their own learning experience.
	Lecture's Energy and Engagement	**9.** Instructors Need to be Energized and Engaged	Students' perceptions of the energy and engagement of the lecturer have a significant influence on how well and how long they can pay attention.
	Self-Identification as Visual Learners	**10.** Visual Aids Improve Learning Experience	Self-identifying as visual learners correlated with belief that they learn the best when the materials are more visual.

5.1 RQ1: What Factors in Online Learning Environments Affect *Learning Gains* (i.e., Measured Difference Between the Post-test and the Pre-test Scores) for Undergraduate Engineering Students?

In this section, I report the answers I obtained quantitatively for the question, "What factors (e.g. types of intervention, presentation of content, activities) in online learning environments affect learning gains, which are measured via post-test scores minus pre-test scores, for undergraduate engineering students?" For each trial, a stepwise multiple regression analysis was conducted to determine the degree to which variables based on the data collected were able to predict students' learning gains. I also present relevant descriptive statistics for each trial. As mentioned in the methods section, the three trials were not combined because topics in each trial are not mutually independent of one another. For reasons that will be explained further below, in order to test additional hypotheses, I then used Tukey's test to compare learning gain differences among the factors.

5.1.1 Multiple Linear Regression on Learning Gain Scores

5.1.1.1 Trial 1

At Step 1 of the analysis, participants' pre-test scores were entered into the regression model and were found to be significantly related to learning gains, $F(1,68) = 124.538$, $p<.001$. This model accounted for approximately 64% of the variance of learning gains (Adj. R2 = 0.642). At Step 2 of the analysis, the writing intervention performance binary value were entered into the model and were also found to be significantly related to learning gains, $F(2,67) = 71.247$, $p<.001$. This model accounted for approximately 67% of the variance of learning gains (Adj. $R^2 = 0.671$). Finally, at Step 3 of the analysis, the level of student's frustration with the intervention was entered into the regression model and was found to be significantly related to learning gains, F (3,66) = 58.725, $p < .001$. This model accounted for approximately 72% of the variance of learning gains (Adj. $R^2 = 0.715$). Figure 5-a below is the output of the SPSS.

Trial 1 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.804[a]	0.647	0.642	1.231
2	.825[b]	0.680	0.671	1.180
3	.853[c]	0.727	0.715	1.098

a. Predictors: (Constant), Pre-Test

b. Predictors: (Constant), Pre-Test, Writing Inter

c. Predictors: (Constant), Pre-Test, Writing Inter, Level of Frust treatment 1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	188.724	1	188.724	124.538	.000[b]
	Residual	103.047	68	1.515		
	Total	291.771	69			
2	Regression	198.457	2	99.229	71.247	.000[c]
	Residual	93.314	67	1.393		
	Total	291.771	69			
3	Regression	212.255	3	70.752	58.725	.000[d]
	Residual	79.516	66	1.205		
	Total	291.771	69			

a. Dependent Variable: Learning Gain

b. Predictors: (Constant), Pre-Test

c. Predictors: (Constant), Pre-Test, Writing Inter

d. Predictors: (Constant), Pre-Test, Writing Inter, Level of Frust treatment 1

Figure 5-a: Output of The SPSS for Trial 1

On the basis of these results, we can see that the learning gains in trial 1were primarily predicted by the pre-test score (PRT), performance of the writing intervention (WI), and the level of frustration with the intervention (FL). Standardized coefficient betas of these variables are presented in Table 5-3 below.

Table 5-3: Summary of Standardized Coefficient Betas

Variable	Variable Type	Standardized Coefficient Beta
Pre-Test Score (PRT)	Continuous	-0.827
Performance of Writing Intervention (WI)	Binary (Yes/No)	-0.235
Level of Frustration with Intervention (FL)	Continuous	-0.225

Figure 5-b and Figure 5-c represent the distribution of the pre-test and post-test scores for Trial 1.

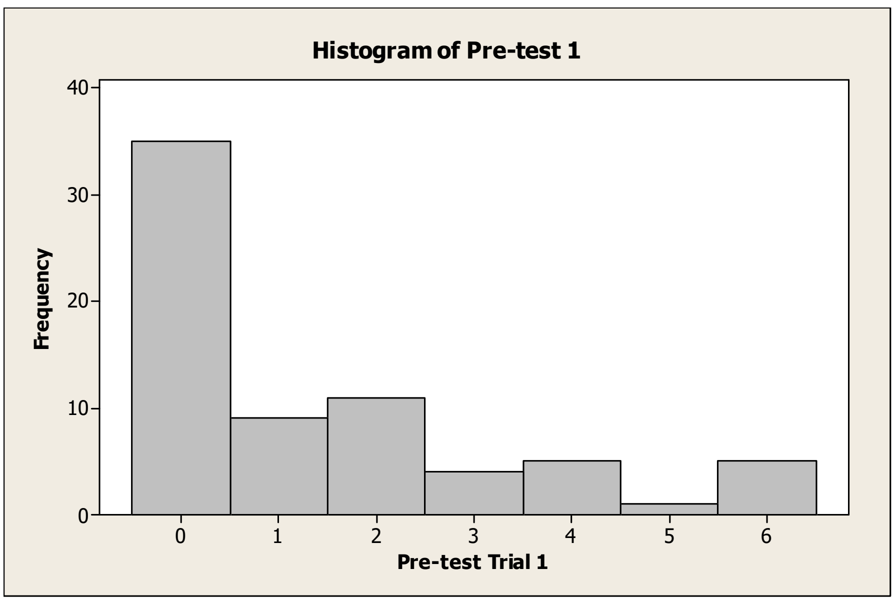


Figure 5-b: Pre-test Score Distribution for Trial 1 (N=70)

Figure 5-c: Post-Test Score Distribution for Trial 1 (N=70)

The average rating of students' level of frustration with the intervention for each of four interventions (A, V, T, V+T) in Trial 1 and their standard deviations are shown in Table 5-4.

Table 5-4: Frustration Level Averages by Intervention Type

Intervention Type	Avg. of Level of Frustration	Std. Dev.
Audio	2.11	1.02
Video	1.89	1.02
Text	1.82	0.81
Video & Text	1.76	0.75

I also reported the difference in learning experience among the four interventions for all trials, which can be seen in Figure 5-d[7]

[7] I was curious about overall patterns of frustrations and so I combined all 3 trials to see the general experience of frustration produced by the type of intervention that they received. The figure shows histograms of all four interventions separated by type. It can be seen that even though the means are approximately 2 for all interventions, the shapes of the tails differ slightly from type to type. The implications will be discussed in Chapter 6.

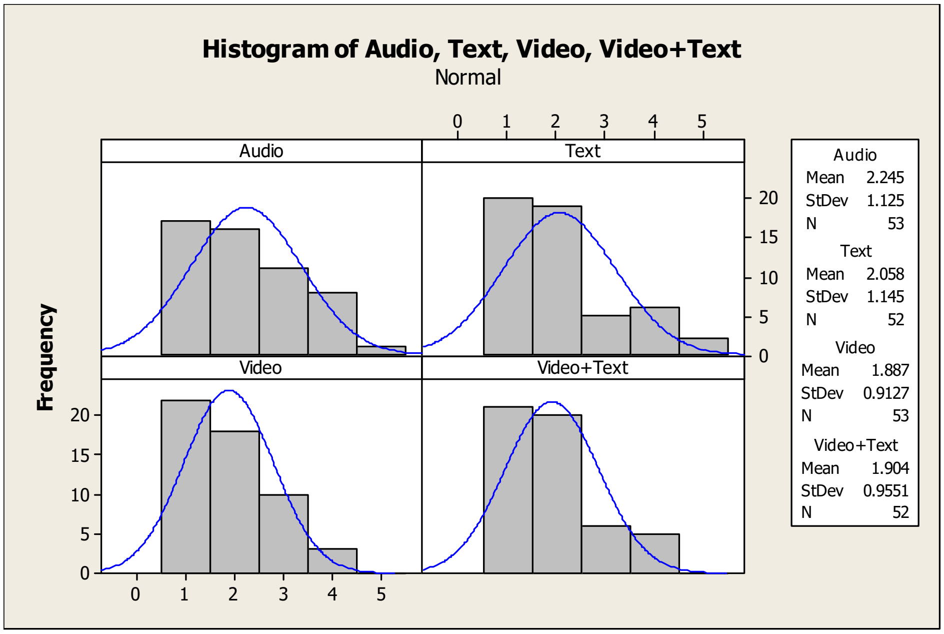


Figure 5-d: Levels of Frustrations of Four Interventions

5.1.1.2 Trial 2

At step 1 of the analysis (and the only step for this trial), the pre-test score was entered into the regression model and was found to be significantly related to learning gains, $F(1,68) = 58.725$, $p < .001$. This model accounted for approximately 59% of the variance of learning gains (Adj. $R^2 = 0.586$). On the basis of these results, we can see that the learning gains in Trial 2 were primarily predicted only by the pre-test score (PRT). The standardized coefficient beta of this variable is -0.769. Figure 5-e below is the output of the SPSS for trial 2.

Trial 2 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.769[a]	0.592	0.586	1.188

a. Predictors: (Constant), Pre-Test 2

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	139.361	1	139.361	98.703	.000[b]
	Residual	96.010	68	1.412		
	Total	235.371	69			

a. Dependent Variable: Learning Gain

b. Predictors: (Constant), Pre-Test 2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.441	0.306		14.491	0.000
	Pre-Test	-0.940	0.095	-0.769	-9.935	0.000

a. Dependent Variable: Learning Gain

b. Predictors: (Constant), Pre-Test 2

Figure 5-e: Output of The SPSS for Trial 2

Figure 5-f and Figure 5-g represent pre-test and post-test scores for Trial 2. This difference in distribution reflects that knowledge gained in Topic 1 is applied to Topic 2.

Figure 5-f: Pre-test Score Distribution for Trial 2 (N=70)

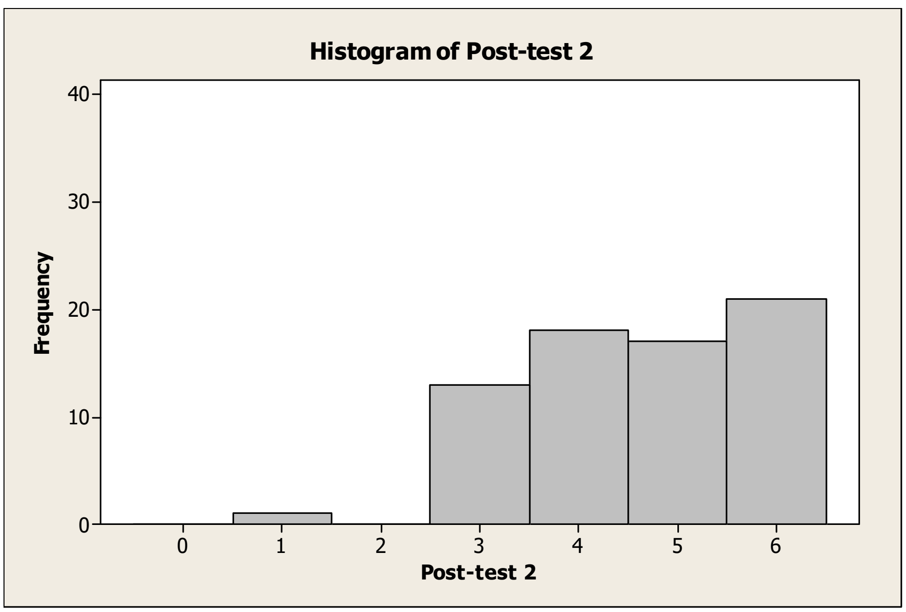


Figure 5-g: Post-Test Score Distribution for Trial 2 (N=70)

5.1.1.3 Trial 3

At Step 1 of the analysis (and the last for this trial), the pre-test score was entered into the regression model and was found to be significantly related to learning gains, F (1,68) = 30.356, p < .001. This model accounted for approximately 30% of the variance of learning gains (Adj. R2 = .298). On the basis of these results, we can see that the learning gains in Trial 3 were also primarily predicted only by the pre-test score (PRT). The standardized coefficient beta of this variable is -0.556. The output of the SPSS for Trial 3 is shown in Figure 5-h.

Trial 3 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.556[a]	0.309	0.298	1.017

a. Predictors: (Constant), Pre-Test 3

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31.410	1	31.410	30.356	.000[b]
	Residual	70.361	68	1.035		
	Total	101.771	69			

a. Dependent Variable: Learning Gain

b. Predictors: (Constant), Pre-Test 3

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.140	0.397		7.908	0.000
	Pre-Test	-0.607	0.110	-0.556	-5.510	0.000

a. Dependent Variable: Learning Gain

b. Predictors: (Constant), Pre-Test 3

Figure 5-h: Output of The SPSS for Trial 3

Figure 5-i and Figure 5-j represent pre-test and post-test scores for Trial 3.

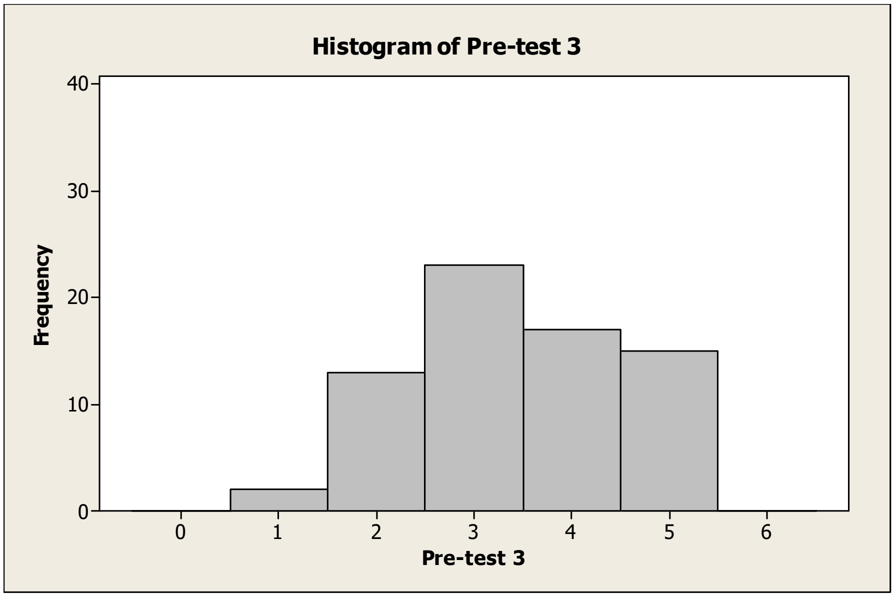


Figure 5-i: Pre-test Score Distribution for Trial 3 (N=70)

Figure 5-j: Post-Test Score Distribution for Trial 3 (N=70)

In summary, the stepwise multiple regression analysis showed that in Trial 1, the main predictors for students' learning gains were the pre-test scores, the performance of the writing reflection intervention, and the level of frustration with the intervention about Topic 1. In Trials 2 and 3, the only predictor for students' learning gains was the pre-test score for each topic. There was no significant relationship between the learning gains and either the performance of the writing reflection or the level of frustration with the intervention for Trial 2 and Trial 3. The implications of these results will be discussed in the following chapter.

5.1.2 Tukey's Test on Learning Gain

Initially, I assumed that the four interventions would be significant factors and, on the basis of this assumption, hypothesized that the learning gain scores would be indifferent. My reasoning was as follows: because the four interventions contained the same information, the learning gains that they produced should have been the same. The way the information was delivered differed, however, so it seemed logical that the students' experiences would differ.

However, when I conducted multiple regression analysis tests to see what factors affected the learning gains and how much each of those factors affected the learning gains, none of the four interventions showed up as a factor; instead, the levels of frustration with the intervention affected students' learning gains.

Since multiple regression analyses show that the first hypothesis (that interventions affected student learning indifferently) was partly true, in that levels of frustration with the interventions were affecting students' learning, I just tested the second hypothesis using Tukey's test to determine whether the second hypothesis holds or does not. The discussion that follows will thus take a slightly different path from the usual. The following is the hypotheses made:

$$H_0 = \mu_A = \mu_T = \mu_V = \mu_M \text{ (all means of leaning gain are equal)}$$

$$H_1 = At\ least\ one\ mean\ is\ not\ equal$$

Figure 5-k shows the Tukey's test results:

Multiple Comparisons						
Dependent Variable:						
Tukey HSD						
		Mean			Interval	
(I) V1		Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Audio	Vid+Text	-0.46	0.686	0.906	-2.27	1.35
	Text	1.01	0.686	0.463	-0.80	2.82
	Video	0.39	0.677	0.939	-1.39	2.17
Vid+Text	Audio	0.46	0.686	0.906	-1.35	2.27
	Text	1.47	0.696	0.160	-0.36	3.31
	Video	0.85	0.686	0.602	-0.96	2.66
Text	Audio	-1.01	0.686	0.463	-2.82	0.80
	Vid+Text	-1.47	0.696	0.160	-3.31	0.36
	Video	-0.62	0.686	0.805	-2.43	1.19
Video	Audio	-0.39	0.677	0.939	-2.17	1.39
	Vid+Text	-0.85	0.686	0.602	-2.66	0.96
	Text	-0.62	0.686	0.805	-1.19	2.43
Based on observed means.						

Figure 5-k: Results of the Tukey's Test showing the learning gains associated with the interventions

We can see in Figure 5-k above that the p values (highlighted with a red rounded rectangle) are higher than the significance value of 0.05. There is no statistically significant difference between the learning gains associated with the interventions when each one is compared with each of the others. We can conclude that it fails to reject the null hypothesis $(\mu_1 = \mu_2 = \mu_3 = \mu_4)$ that there is no difference in between learning gains by interventions.

5.2 RQ2: What Factors in Online Learning Environments Affect the Learning Experience for Undergraduate Engineering Students?

To address this question, I initially used thematic analysis and descriptive statistics to explore what factors were affecting students' learning experience. Qualitative results showed that the factors affecting students' learning experience were the four interventions, the writing interventions, repetition, and inconsistency of terminology. I then used the Tukey's test on the four interventions to quantitatively compare the available and relevant quantitative results (mean levels of frustration).

5.2.1 Qualitative Result: Thematic Analysis

In this section, I report the answers I obtained qualitatively for the research question RQ2: What factors in online learning environments affect the learning experience for undergraduate engineering students? For purposes of the analysis, findings are reported in reference to participants': (1) general learning experience. I report the seven major themes that emerged from the qualitative interview data and descriptive statistics.

In the following sections, I report the answers I obtained quantitatively for the research question RQ2: What factors in online learning environments affect the learning experience for undergraduate engineering students? This section is divided by (1) findings associated with intervention type and (2) findings associated with general (i.e., not associated with any particular intervention) themes.

5.2.1.1 Intervention Type

In this section, I report the findings associated with students' learning experience according to intervention type.

5.2.1.1.1 Audio-only Intervention Theme: Audio-only Intervention Not Visual, Not Helpful

The Audio-only intervention presented to participants was not visual enough and thus it was not helpful to them. Thirty-five of the 53 students who received the Audio-only intervention ranked the intervention as 3 (i.e., liked the least) out of the three interventions they received. Of these thirty-five students, more than half specified why they ranked it as number 3; the rest of the group did not. Many students specified that the Audio-only intervention presented was not helpful to them because it lacked visual elements. This suggests that students felt that the Audio-only intervention was not particularly effective in enhancing their learning experience. Figure 5-l below displays the distribution of the students' ranking of the Audio-only intervention as compared to the other interventions they received (i.e., A,,T, V; A,V, VT-, and A, VT, T-) Text).

```
Stem-and-leaf of Audio  N  = 53
Leaf Unit = 0.10

 4    1  0000
 14   2  00000000000000
(35)  3  00000000000000000000000000000000000
```

Figure 5-l: Students' Ranking of the Audio-only Intervention (1 - highest and 3 - lowest)

Twenty-five of 35 students for whom Audio-only was ranked number 3 identified the problem as the lack of a visual representation, which they said made learning the materials difficult. For example, students said:

> *"The audio was kind of hard to follow because there was nothing visually to go by."*
>
> *"The audio was easier to get distracted because there's nothing to look at."*
>
> *Interviewer: Why did you rank the audio intervention as a number 3?*
> *Participant: "I don't know exactly why, but there's nothing visual with it."*

Ten of the 35 students for whom Audio-only was ranked number 3 indicated (in response to a direct question, as in the quotation above) that the audio only intervention was their last choice, but did not specify why it was not helpful or why they preferred the other interventions.

In summary, most students felt that the audio only intervention helped them the least and overall, negatively affected their learning experience.

5.2.1.1.2 Text-only Intervention Theme: Information Quantity Overwhelming, yet Allowed Self-pacing

The quantity of information in the Text-only intervention presented to students was perceived by some as overwhelming, but other students found it helpful that they could process the information at their own pace. Twenty-seven of the 52 students who received the Text-only intervention ranked it number 3 (the lowest) out of the interventions they received. This ranking indicates that the majority of the students who received the Text -only intervention said that it was their third choice because, as suggested by those students, the quantity of information presented to them was overwhelming. However, unlike the Audio-only intervention, the text intervention was regarded by a number of students as having positive traits. Nine students gave it the ranking 1 and sixteen students gave it the ranking 2; some specified that this was because they could process the information at their own pace. Figure 5-m displays the distribution of how the students ranked the Text-only intervention among the three interventions they received. Representative quotations are below.

```
Stem-and-leaf of Text  N  = 52
Leaf Unit = 0.10

  9    1  000000000
 16    2  0000000000000000
(27)   3  000000000000000000000000000
```

Figure 5-m: Students' Ranking of the Text-only Intervention (1 - highest and 3 - lowest)

Fifteen of the 27 students who gave Text-only a 3 ranking said that the reason for their low ranking was that it provided them with too much information at once. For example, students said:

> *"There was a lot of text. I was thinking 'I don't want to read this.'"*
>
> *"It was just too much text. I was thinking 'there's so much information at once.'"*
>
> *"When I get thrown a wall of text, it's like, 'Just give me the information'. I told you, I learn a lot of my stuff just listening to the podcasts and stuff."*

Twelve of the 27 students indicated in response to a direct question that the Text-only intervention was ranked number 3 but did not specify why they ranked it number 3 or why they preferred the other interventions.

Nine of the 25 students who assigned a ranking of 1 or 2 to the text only intervention gave these rankings because they were able to move through it at their own pace, whereas for the video or audio only, the pace is determined by the media player. For example, students said:

> *"I put the text [intervention] above the video just because [it was helpful] … being able to read about it on my own or with my own pace."*
>
> *"I mean definitely the text [intervention] kind of made the definitions a little more concrete cause you can actually read it at your own pace."*
>
> *"I liked the text [intervention] because you're going to go at your own pace. So especially if it's something you're learning for the first time, it's kind of nice to just go through at your own pace and make sure you understand each second."*

In summary, approximately half of the students felt that the Text-only intervention was not particularly helpful to their learning experience because an overwhelming quantity of information is presented at once. The other half found it helpful in their learning experience because it enabled them to process the information at their own pace.

5.2.1.1.3 Video Intervention Theme: Seeing the Lecturer's Engagement with Materials Helpful

The visual aspect of the Video intervention was perceived as helpful because it enabled students to see the lecturer engage with the material through actions. Eighteen of the 53 students who received the Video intervention ranked it number 1 (the highest) out of the interventions they received. The majority of students indicated that the visual aspect of this intervention helped them learn the best because it enabled them to see the lecturer engage with the material through action (e.g., the lecturer's moving of the cursor, pointing, drawing, etc.). All of these students indicated that the Video intervention was beneficial, and most of them also indicated the way in which the video intervention helped them, see Figure 5-n. Representative quotations are below.

```
Stem-and-leaf of Video  N  = 53
Leaf Unit = 0.10

 18   1  000000000000000000
(31)  2  0000000000000000000000000000000
 4    3  0000
```

Figure 5-n: Students’ Ranking of the Video Intervention (1 - highest) and (3 - lowest)

Fifteen of the 53 students suggested that the video intervention was helpful because it provided them with a visually supported detailed explanation that helped them to better understand the topic. For example, students said:

> *“I feel like in the video there was a more in-depth explanation of the graphs, figures, and the numbers.”*
>
> *“The video was really helpful because they would draw something on the picture, or you could even see the mouse reference where they are and what and where they are indicating.”*
>
> *“It was helpful when the video had the professor drawing right on it. “*
>
> *“He'd be drawing all these different things on the board and making a connection to other parts.”*

In summary, students said that the Video intervention added value to their learning because it permitted them to benefit from the lecturer’s visible interaction with the learning material, which reinforced their memory of what they were learning. These statements indicate that the majority of the students found that the video intervention was helpful and suggest that students felt the video intervention was effective in enhancing their learning experience.

5.2.1.1.4 Video+Text Intervention Theme: Variety in Choice of Interventions Helpful, Provides Supplements

The Video+Text intervention enabled students to choose among various types of learning interventions, and the other types of intervention that they did not choose initially acted as a supplementary intervention. Twenty-six of the 52 students who received the Video+Text intervention indicated that they ranked the intervention number 1 and said that it helped to support their learning (see Figure 5-o). Eighteen of these students also expressed the way in which the mixed intervention was helpful to them. The majority of students said that the mixed intervention helped them because it gave them a choice of intervention and the option of using a different intervention if their first choice did not help them sufficiently. This suggests that the mixed intervention was helpful in improving their learning experience. Representative quotations are below.

```
Stem-and-leaf of Video & Text  N  = 52
Leaf Unit = 0.10

 26  1  00000000000000000000000000
 21  2  000000000000000000000
 5   3  00000
```

Figure 5-o: Students' Ranking of the Video+Text Intervention (1 highest) and (3 lowest)

Ten students said that the mixed intervention was helpful to them because not only did it give them the freedom to choose what type of intervention they wanted to focus on, it also gave them the supplemental intervention simultaneously. For example, students said:

> *"I think that combined [mixed] would be the best, just because having multiple stimulants is better for people who learn differently, just making sure that you're getting all aspects of the best way to learn. Then visual, I [like] more also rather than just reading text, because I feel like it's more engaging. If I'm just reading text, I'm not necessarily connecting [to] what I'm reading."*

> *"I learn better from reading rather than listening, so having the text there along with the video was really helpful because you can read the script and see what they were doing."*

> *"Having the text there along with the video is really helpful because you could read and see what they were doing. Just the video was a little bit hard to just listen and look at the graphs."*

Eight students stated that the mixed intervention was beneficial because if they did not understand the material from the first intervention they looked at, then they could go to the second one to get an additional explanation. For example, students said:

> *"For a more difficult and not necessarily intuitive concept, having the text [intervention in addition to a video] was beneficial."*
>
> *"I watched the video but that was confusing, I just read through those texts and just seeing it helped."*
>
> *"It was nice to have them both [the video and the text intervention] in case there was something I didn't quite get."*

Eight students indicated that the Video+Text intervention was their top-ranked choice but did not specify why it was the most helpful or why they preferred it to the other interventions.

In summary, the majority of students felt that the Video+Text intervention enhanced their learning experience because it allowed them to choose between two types of learning methods presented simultaneously and to use the second intervention if the first intervention did not help them enough.

An additional point of interest is the response of the thirty-five students who had both the Video intervention and the Video+Text intervention. Sixteen of these 35 students indicated that they preferred the intervention with just a video because they felt that the additional intervention was excessive information. For example, students said:

> *"Video only would be a 10. And then the video with text [a] 9. Because there was more to do with the video and text. More to read up there."*
>
> *"[Rank]1, I would say I prefer video... I would give a [rank] 3 to the video and the text because I would say because you have to look at the video and listen to it and try and read that it's a little too distracting, so I'd give that a 3."*
>
> *"Video and text would be [Rank] 2. They [Text] were just kind of confusing to me."*

For students who had both the Video and Video+Text intervention, approximately half of the students indicated that they preferred Video alone because of the superfluous information supplied by the text.

5.2.1.1.5 Writing Reflection Intervention Theme: Writing in Own Words Useful and Engaging

Writing in their own words (i.e., paraphrasing and summarizing) supports learning through engagement. All 35 students who had the writing intervention expressed some opinions regarding how and why the writing reflection intervention affected their learning. Most of the students reported that the writing reflection helped them with their learning, although a small minority said that it did not. Half of the students specified why it helped; the other half did not. Representative quotations are shown below.

Eight students said that the reason the writing reflection intervention was helpful was that they had to paraphrase and summarize during the writing reflection. They stated that it was beneficial to write in their own words because it helped with their memory. For example, they said:

> *"This [writing reflection] added value because I think that while I was doing writing reflection, I wrote things in my own words because I'm not exactly writing down, recording exactly what the lecturer said."*
>
> *"I think it helps. Whenever I write something down in my own words, I remember it more."*
>
> *"I think it was definitely helpful because it kind of helped me to summarize each one in my own words from what I remembered from it. I think definitely writing helped memory"*

Seven students indicated that the reason that they found it helpful was that the expectation of having to do a writing reflection made them more engaged in learning. In support of this, students said:

"I would say it benefited [me] when you asked me to do that [writing reflection] because then I had to be like, 'Oh, he's expecting me to reflect.'"

"I mean, it [the writing reflection] forces you to engage with the material a bit more. So ... it's positive."

"The writing reflection is a really good tool. You pick the thing, and you write it down... So, personally, I feel good about writing some good points, or main points, or useful points about what I'm learning."

Fifteen students indicated in response to a direct question that the reflection was helpful but did not specify why it was helpful.

In contrast, five of the 35 students who did the writing reflection indicated that they felt that it did not help to increase their learning or helped minimally. One of the five students said that the writing reflection had a minimal effect because the topic was not complicated enough that paraphrasing added value to the learning and another student stated that writing alone was not enough. The remaining three did not specify a reason.

In summary, most students felt that the writing reflection helped them to better understand and remember the material.

5.2.1.2 Factors from General Findings

In this section, I report the most emergent themes from the findings that are associated with students' learning experiences that are not tied to any particular intervention type.

5.2.1.2.1 Inconsistency of Terminology Theme: Efforts to Understand Terminology Helpful for Understanding

Students said that putting in the effort to figure out the connections or relationship between discrepant terms helped them understand the material. Drawing from the pilot study presented in Chapter 3, I searched throughout the codebook and interview data for factors that can produce value-added frustration (material that makes students more engaged and thus leads to the strengthening of their understanding or better long-term retention of the materials).

One factor that produced frustration that may have added value to their learning was inconsistency in terminology (e.g., between the primary video and the intervention or the content quiz questions and the intervention). Ten students expressed frustration with the inconsistency of terminology. However, some of them also suggested that putting in the effort to figure out the connections or relationships between discrepant terms actually helped them understand the material. For example, students said:

> *"I guess one thing I was frustrated with was the questions used different terminology than in the video. it was frustrating to have to figure out those."*
>
> *"I ended up getting it, but I don't think it explicitly uses the terminology that the questions used."*
>
> *"I think some of the terminologies were confusing, the videos used a lot of different terms [from the primary video]."*

In summary, we can see that inconsistent terminology created confusion and thus may have initially detracted from students' learning. Some students remarked that their effort to make sense of the discrepancy ultimately helped them learn better; thus, there appears to be a slight chance that inconsistency of terminology can be seen as an example of desirable difficulty; see the additional discussion in the following chapter.

5.2.1.2.2 Repetition Theme: Repetitious Content Annoying and Helpful

Some students reported that the repetition of key materials in the online learning environment created annoyance, but others judged it to be helpful. Twenty-four of 70 students indicated that they felt that the repetition of materials in quizzes and the overall design of the environment affected their learning. Some students stated that repetition in the online learning environment created annoyance, but also some expressed the opinion that it was helpful. Representative quotations are below.

Fourteen of 24 students said that redundancy was helpful because of the similarity in the quiz or that the topic reinforced what they learned. For example, they said:

> *"Seeing it over and over again, it helped me. That helped a lot actually."*
>
> *"I liked that there were quizzes and they kept coming up with similar questions over and over because that helped me see it again"*
>
> *"I liked how some of the information was repeated on both of the videos. It reinforced what I thought it was."*
>
> *"All the questions that were first asked I answered I don't know because it was new. And then after the first round, I got some ideas about the knowledge. And like by the time it got to the third video, I was kind of expecting what's coming."*
>
> *"Because of the repetition, it essentially, as it repeated more I had to spend less and less time trying to learn it. And more and more time just like feeling confident that I understand what it's saying."*

In contrast, ten of 24 students indicated that they felt that redundancy negatively affected their learning experience. The repeating of questions or topics created annoyance. Students stated:

> *"Why am I doing this repeatedly?... I felt like it was unnecessary. But, I mean, I'm an engineering student, so maybe other people would... How he's explaining the relation between X and Y on an exponential, I felt like he spent a little more time, so I felt like I could just not watch it."*
>
> *"I disliked the repetitiveness. so, risk-neutral first, and then it went to adverse then it went to seeking... So maybe I just [wanted] all at once."*
>
> *"I almost felt too similar to the one before. It was just getting like frustrating"*

To summarize, repetition was seen as helpful by some because the similarity in materials in quizzes and overall design of the environment strengthened what they learned and reduced the effort they had to make to learn it; in contrast; repetition was seen by others as unnecessary and created annoyance for students who thought that the topic was too simple to warrant repetition. The following Figure 5-p shows the distribution of GPA for students who stated that the experience of the repetition was positive or negative. The implications will be discussed in Chapter 6.

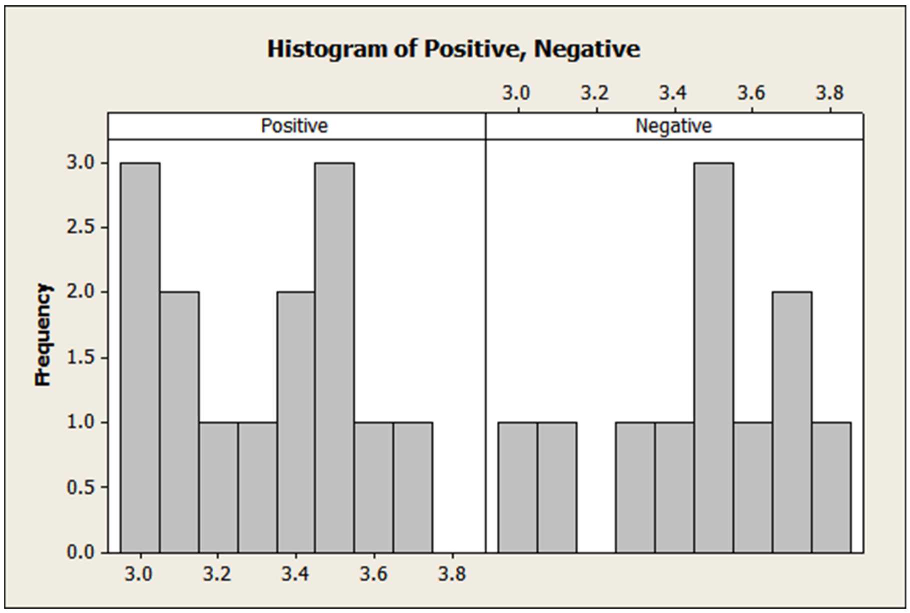


Figure 5-p: Histogram of GPAs: Positive and Negative Experience (n=24)

5.2.2 ***Quantitative Result: Tukey's Test***

I used Tukey's tests to find statistical differences between mean experience (levels of frustration with interventions) across the four interventions. The following is the hypothesis made:

$$H_0 = \mu_A = \mu_T = \mu_V = \mu_M \text{ (all means of level of frustration are equal)}$$

$$H_1 = At\ least\ one\ mean\ is\ not\ equal$$

Figure 5-q shows the Tukey test results:

Multiple Comparisons						
Dependent Variable:						
Tukey HSD						
		Mean			Interval	
		Difference			Lower	Upper
(I) V1		(I-J)	Std. Error	Sig.	Bound	Bound
Audio	Vid+Text	0.34	0.203	0.335	-0.18	0.87
	Text	0.19	0.203	0.792	-0.34	0.71
	Video	0.36	0.202	0.288	-0.16	0.88
Vid+Text	Audio	-0.34	0.203	0.335	-0.87	0.18
	Text	-0.15	0.204	0.875	-0.68	0.37
	Video	0.02	0.203	1.000	-0.51	0.54
Text	Audio	-0.19	0.203	0.792	-0.71	0.34
	Vid+Text	0.15	0.204	0.875	-0.37	0.68
	Video	0.17	0.203	0.834	-0.35	0.70
Video	Audio	-0.36	0.202	0.288	-0.88	0.16
	Vid+Text	-0.02	0.203	1.000	-0.54	0.51
	Text	-0.17	0.203	0.834	-0.70	0.35
Based on observed means.						

Figure 5-q: Tukey's Test Levels of Frustration With Interventions

We can see in Figure 5-q above that the p-values (highlighted with a red rounded rectangle) are higher than the significance value of 0.05. There is no statistically significant difference between the levels of frustration associated with the interventions when each one is compared with each of the others. We can conclude that it fails to reject the null hypothesis that there is no difference in the levels of frustration associated each of the interventions.

The incongruence between quantitative data and the qualitative data will be thoroughly explored in the discussion section.

5.3 RQ3: What Factors in Online Learning Environments Affect Undergraduate Engineering Students' Self-Reported Memory?

In this section, I report the answers I obtained qualitatively for the research questions R3: What factors in online learning environments affect the self-reported memory for undergraduate engineering students? For purposes of the analysis, findings are reported in reference to participants' self-reported memory. I report the three major themes that emerged from the qualitative interview data and descriptive statistics.

To investigate whether there were factors that helped students' self-reported memory, I initially identified themes that arose from a question about their best lecture experience. For the students, I defined this as "the best lecture or teacher that they had ever had.' I asked them to think particularly about cases where they still remember the class materials. I reasoned that if students still remembered the lecture material at this point in their academic careers, then I could assume that the lecturer that they were thinking of may very likely have done a good job of helping the student learn the material. Thus, I wanted to identify the salient qualities of those lectures or lecturers and determine whether they aligned with the qualities found in the online learning materials examined in this study.

Sixty students reported that they had had a "best lecture/lecturer experience." Forty-five of those 60 students indicated that they remember that the professors or classes were engaging when professors were passionate about the topic and energetic when they taught it and/or when the lecturers used multiple visual aids. Using this information, I searched the codebook for instances of these themes in the data. Three emergent themes about factors that influenced participants' self-reported memory were (1) participation in interactive tasks, (2) the energy and engagement displayed by the instructor, and (3) self-identification as a visual learner (although instances of the theme related to visual aids do occur, the students do not explicitly link them to memory). I discuss each of these themes in the following sections.

5.3.1 Interactive Tasks: Interactive Task Participation Supports Learning, Confidence

Participation in interactive tasks positively influenced students' learning and made them more confident about their own learning experience. To understand how participating in interactive tasks influenced students' self-reported memory, I examined occurrences in the codebook and interview data in which students discussed performing their interactive activities.

Students talked about two interactive elements in this study: (1) the writing reflection intervention and (2) the content quizzes. The first of these, the writing reflection intervention, primarily addressed Research Question 1, but students did say also that the writing intervention helped them remember more and engage more with the material; please refer to section 5.2.1.1.5. This indicates that the writing reflection intervention helped them with self-reported memory to some degree.

The second interactive element was the content quizzes. Students expressed that the action of figuring out the answers in the content quizzes helped them remember more and the quizzes themselves acted as checkpoints to show how much they recalled, which gave them an increase in self-efficacy (confidence in their understanding of the materials). Specifically, 25 of 70 students expressed in their overall experience that they felt that content quizzes helped to enhance their memory. Some students also explained different ways in which they were helpful. Representative quotations are below.

Eleven students said that the quizzes were helpful because they required mental effort that produced engagement and acted as checkpoints where students stopped to reflect on how much they remembered and assess their sense of how much they understood. For example, students said:

> *"I think those [quizzes] are very helpful because testing your knowledge, it makes you realize 'Oh, I probably understood something,'"*
>
> *"I liked how it [quizzes] came back to the same questions and it tells you a little bit more about them and then it would ask these questions."*
>
> *"In general, I liked the quizzes. Those were a good way to test your knowledge."*

In summary, students repeatedly indicated that the content quizzes made them more confident about their own learning experience and also served as checkpoints to test their knowledge. Using the data from the participants' discussion of both the writing reflection and the interactive quizzes, I found that students indicated that participation in interactive tasks influenced their learning.

5.3.2 Lecture's Energy and Engagement Theme: Instructors Need to be Energized and Engaged

Students' perceptions of the energy and engagement of the lecturer have a significant influence on how well and how long they can pay attention; studies show that high levels of engagement and energy increase students' attention span. Half of the students said that the primary lecturer's low energy and/or the length of the primary lecture video caused a lack of attention or reduced their participation in the study. In particular, 39 of 70 students expressed in response to a question about their overall experience that they felt that the lecturer's low energy and the long duration of the main video made it difficult for them to pay attention. It is noteworthy that students' perception of the length of the video was that it was 'too long', even though the longest of the videos was 6 minutes, which suggests that the lecturer was not engaging. Some students gave more detailed explanations of how these factors influenced their attention, which is associated with how much they remember. Representative quotations are below.

Eighteen students said that the primary lecturer's low energy and lack of enthusiasm made it difficult for them to pay attention. For example, students cited the speaker's monotone voice, lack of liveliness, and low levels of enthusiasm as examples of why they lost interest in the video. For example, participants said:

> *"And like a monotone voice. It was difficult to pay attention to."*

"Nothing against the video, but I felt it could be a little bit more lively if that makes sense."

"I think it could be a correlation but not causation, but I would say this teacher's enthusiasm was pretty low. I would say teacher enthusiasm contributes a ton to whether or not students pay attention."

Fifteen students stated that the length (3 to 6 minutes) of the videos was a problem because this span of time was too long for them to pay attention. For example, participants said:

Interviewer: "During the experiment, did you have any time where you could not pay attention? Participant: "Yeah, some of the videos seemed too long."

"I think I wished [video] turned faster because I got a feel of what was going on. I have a really short attention span."

"I was fidgeting a bit but one of the videos was I think six minutes. So, that was one of the longer ones…Because I always had attention issues."

In summary, students got tired of the learning material when they perceived the teacher's energy as too low and/or the three to six-minute videos as too long. These findings show that students' perceptions of the energy and engagement of the lecturer have a significant influence on how well and how long they can pay attention. The reason that the videos felt too long, at least in some cases, is that the lecturer was not engaged enough. Overall, the findings show that students' perception of the energy level of the lecturer affects students' perception of whether or not videos are too long and also their ability to maintain their attention for as long as they need to.

5.3.3 Self-Identification as Visual Learners Theme: Visual Aids Improve Learning Experience

Students' self-identification as visual learners correlated with their belief that they learn best when the materials are more visual. The majority of students stated that one of the factors that made them remember materials the most in their past lectures was the lecturer's use of multiple visual aids; therefore, I decided to look into instances in the codebook related to self-reported memory. Although the codebook and the interview data do show numerous instances of references to the way in which visual aspects of instructional material enhanced their learning experience, students typically did not associate visual aids explicitly with self-reported memory in this online learning experience. As shown in the previous results about interventions, the majority of the students stated that they found interventions with more visuals more helpful and suggested that they were effective in enhancing their learning experience. Twelve students stated that they self-identified as a visual learner, and these students associated visual elements with a good learning experience. They said:

> *"They [video and mixed intervention] both kind of were the same. But I mean, I'm a visual learner, so that video definitely helps. But being able to read too. So I'd say video and text."*

> *"I did [rank] that one [Video+Text] the best because I'm a visual person. A visual person in the sense for having the graph but also reading ... Hearing something, I only remember so much"*

> *"When it was the one that was just text, I sort of just glossed over that. If there is something as a user that I can input into my profile, I'm a visual learner and therefore I want "x" when I need visual help."*

In summary, students stated that they self-identify as visual learners and believe they learn best when the materials are more visual. Implications of all these findings are discussed in greater depth in Chapter 6.

Chapter 6 Discussion

In this chapter, I present multiple takeaways and insights from this research. First, I will briefly recap the general overview of what was expected and what was found. I discuss the features of interventions that affected students' learning and I propose corresponding recommendations for the design of online learning that take advantage of the factors that were shown to improve learning. Lastly, I discuss the role of other factors from the research findings that affected students' learning (e.g., repetition and the energy of the lecturer), and I interpret the findings.

6.1 General Recap of the Research, Hypotheses & Results

To briefly recap the general research approach, my research questions are:

1. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning environments affect *learning gains* (i.e., measured difference between the post-test and the pre-test scores) for undergraduate engineering students?
2. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning environments affect the *learning experience* for undergraduate engineering students, and, specifically, what factors produce desirable difficulty?
3. What factors (e.g., demographic information, intervention type, delivery of information, pedagogical approach) in online learning affect undergraduate engineering students' *self-reported memory*?

My corresponding hypotheses for this research were:

- *Hypothesis #1:* Learning Gains

- *Hypothesis #1A: I hypothesized that* interventions (supplementary information) in various forms (four different types) *would be significant. Making the assumption that the first part is true*, I then hypothesized that students' learning gains scores would not differ by type of intervention.
- *Hypothesis #1B:* I hypothesized that if students performed the writing reflection, then their learning gains would be higher than the learning gains of those who did not.

- *Hypothesis #2:* Learning Experience
    - *Hypothesis #2A:* I hypothesized that if students received new information in the various forms in which interventions were delivered via their auditory and visual channels, then their experience would be differentially affected according to the types of interventions received, specifically:
        - The Audio-only intervention would be the least beneficial to the students
        - The Text-only intervention would be the third most beneficial
        - The Video intervention would be the second most beneficial
        - The Video+Text intervention would be the most beneficial.
    - *Hypothesis #2B:* I hypothesized that if students perform the writing reflection task, they would benefit from the experience.
- *Hypothesis #3: Self-Reported Memory*
    - *Hypothesis #3A:* I hypothesized that if students received new information in the various forms in which interventions are delivered (supplementary

information), then their self-reported memory would be differentially affected according to the types of interventions received, specifically:

    - The Audio-only intervention would be the least beneficial to the students.
    - The Text-only intervention would be the third most beneficial.
    - The Video intervention would be the second most beneficial.
    - The Video+Text intervention would be the most beneficial.

  - *Hypothesis #3B:* If students perform the writing reflection, then students would benefit from the experience.

I tested these hypotheses using both quantitative and qualitative methods to examine the experiences of seventy undergraduate engineering students. The following are the expected and unexpected results:

- ***Intervention, Frustration Level Affects Learning Gains and Experience.*** As expected, the quantitative data showed that the forms in which the students received the new information (i.e. Audio-only, Text-only, Video, Video+Text) affected students' learning gains and experience. However, unexpectedly, the determining factor was not the form itself; more important for the results was the magnitude of the student's reported level of frustration (measured on a Likert scale) with the particular interventions they received. That is, the level of frustration had more of an effect than the intervention form.
- ***Writing Reflection Significant, Negatively Correlated to Learning Gains.*** Also as expected, in the quantitative results, the performance of the writing reflection was significant, but unexpectedly, it was negatively correlated with the learning gains. This

is surprising because, in the interview data, many students expressed that the performance of the writing reflection was helpful in their learning.

- ***Other Factors' Influence on Learning Experiences.*** Unexpectedly, analysis of the qualitative data in light of the quantitative results revealed that other factors that had not been intuitive for me to measure at the time of the experiment also affected the learning experiences significantly (e.g., the instructor's vocal tone, inconsistency in terminology, the presence or absence of visual aids, repetition).
- ***Variance in Desirable Difficulty Across Intervention Types.*** As expected, there is evidence that the Text-only intervention did produce desirable difficulty. However, the Audio-only intervention apparently did not. The writing reflection intervention also produced some desirable difficulty, which might account for the unexpected negative correlation with learning gains, as will be further explained below. Repetition and inconsistency of terminology can be conjectured to have contributed to frustration in some instances but did produce desirable difficulty in some other cases.
- I expected that some interventions would affect self-reported memory, but unexpectedly the qualitative results revealed that interactions, the energy of the teacher and self-efficacy of visual learning determined self-reported memory were largely the determinants of what students said about their memory of the material.

6.2 Connection of Findings to Existing Theoretical Framework

The results are not straightforward and the story they tell is not obvious. In an effort to connect all the dots, after carefully examining the qualitative and quantitative results, both separately and together, I observed that there is one thing common to all the important factors affecting learning gains and the learning experience: frustration. The differences among effects of the factors (both interventions and other factors) may come from how much a given factor increases or reduces the level of frustration and whether or not learners are able to overcome the frustration, which is the essential thing that must happen if they are to be able to achieve learning gains and an improved learning experience.

To elucidate and explain these abstract findings concretely within the framework of a theory, I revisited the literature on cognitive load theory to discover possible connections between frustration and the salient features of the theory. Cognitive load theory states that our mental capacity (working memory) is limited and that information is processed in working memory by two channels, audio and visual. Please keep in mind that the term *information*, in this theory, is equivalent to the load (i.e., the knowledge or task) that learners put cognitive effort into processing. Because the terminology used in cognitive load theory is central to the discussion that will follow, it is essential to understand its usage. Also in the theory, there are three types of load that the learner's mental capacity must handle: (1) intrinsic cognitive load (ICL), (2) extraneous cognitive load (ECL), and (3) germane cognitive load (GCL). Intrinsic cognitive load is the inherent complexity of the learning task and material; (e.g., quantum physics has a high relative inherent difficulty, a simple addition task (e.g. 1+1=2) has low relative inherent difficulty). Extraneous cognitive load can be best understood as the way information or tasks are presented to a learner, that is, "information" that does not have a payoff in learning (e.g., dealing with background noise; trying to understand a lecturer's unreadable handwriting). Germane load is the cognitive work necessary to connect incoming information to existing knowledge (e.g., when a learner is able to understand a new topic such as risk-seeking through associating it with a familiar context of risk-neutral behavior).

Most educational researchers agree that the objective for the designers of educational material is, according to the cognitive load theory, to minimize the extraneous load and to maximize the germane load for a better learning outcome (Schweppe & Rummer, 2016). From the literature, I initially concluded that the extraneous load is equivalent to “bad” frustration and that eliminating these bad frustrations would make the students’ learning better. However, this oversimplified conclusion appears to explain only half of the story; it did not explain my findings on the positive role of frustration in learning.

I found in other research studies the seemingly contrary and counterintuitive concept called “desirable difficulty,” which states that making learning more difficult (demanding of effort) leads to desirable learning outcomes, as indeed some of my findings suggest (Bjork, 1994). To see how these two contrary recommendations for learning (i.e., minimize difficulty that results in frustration vs. provide difficulty that increases frustration) can both produce good results, let us also view desirable difficulty through the lens of cognitive load theory.

According to Roediger III and Karpicke (2006), effects attributed to desirable difficulties are stronger and often observable only when performance is tested after a delay. In terms of cognitive load theory, desirable difficulty is the requisite mental effort put forth by a student to process both intrinsic difficulty of the information and extraneous difficulty that is neither too low nor too high and store the learned material in long-term memory. Therefore, I posit that, for the purposes of this discussion, the intrinsic load can be defined as the intrinsic difficulty of information and extraneous load as external difficulty imposed on the information.

In this study, I intended that the intrinsic difficulty of information would remain constant within each trial so that I could identify the effects of varying the external factors (e.g., type of intervention, teachers' presentation of information, etc.). Across the three trials, the extraneous (external) difficulty remains constant and the intrinsic difficulty varies as a result of students' increasing familiarity with the topic, as will be discussed further. The effect of the intrinsic difficulty of the information was inferred from learners' pretest scores and effects of external difficulty of the information delivery were inferred using the level of frustration. Furthermore, various data (such as the results from the thematic analysis) were used in combination with these indicators to show the effects of the changing intrinsic difficulty and external difficulty of information. In the students' comments and quantitative data, frustration was an important emotional state in their overall perception of the difficulty of information and learning experience. Quantitative and qualitative data were used together to explain students' level of frustration.

6.3 Intervention Features and Effects on Student Learning

I find that there are several possible explanations that might account for the diminishing effects of interventions across the trials. The effects of interventions may not seem positive if assessed strictly from the quantitative perspective. However, if we look at qualitative data, it shows that some interventions do seem to add value to students' learning; thus, we need to look at both closely to determine what is occurring. Before proceeding, however, one caveat is necessary: in the explanation of the results, I will mainly focus on the learning gain score with respect to the results that emerged from Trial 1 because, as will be covered in detail in the next section, the Trial 2 and Trial 3 learning gain scores were significantly affected by other factors (students' increasing ability to infer and the effect of repetition).

6.3.1 Examining the Diminishing Effect of Interventions: Differences between Trial 1 and Trials 2 and 3

I assumed that the five interventions were going to be critical factors affecting students' learning gains in all three trials and that different intervention types might have different effects on learning experiences. Although I did not incorporate it into the hypothesis proposed, I also assumed that the effects of the interventions might decrease slightly over the course of the three trials. These hypotheses were partially confirmed; first, interventions were significant for Trial 1. Further, the effects of different interventions did differ, and the effects were shown to decrease dramatically from the first trial to the second and third. Unexpectedly, however, some hypotheses were not supported. In fact, no interventions were significant for Trials 2 and 3. For Trial 1, the quantitative results show that the students' learning gains were significantly negatively correlated to the level of frustration with the four interventions (Audio-only (A), Video (V), Text- only (T), Video+Text (VT)) and the performance of the writing intervention, However, in the second and third trials, the level of frustration with the four interventions (A, V, T, VT) and the performance of the writing intervention were no longer significantly correlated to students' learning gains. There are two possible explanations for these counterintuitive results: (1) students' increasing ability to infer and (2) the effect of repetition.

6.3.1.1 Students' Increasing Ability to Infer

The first explanation is that since the knowledge of the topics that was presented was building throughout, students' ability to make inferences about future topics diluted the effect of interventions. For example, after students learned about the risk-neutral topic, they were able to infer something about the topic of risk-seeking, and then, with more information, they were able to infer more about the risk-averse topic. This explanation is likely because pre-test scores changed radically as the experiment progressed from the first trial to the second and third. As shown in Figure 5-b in Section 5, the concentration of pre-test scores was near 0 for Trial 1.

This was before students had learned anything about the topic; therefore, I can assume that the pre-test task was perceived as just too hard to relate to their previous knowledge schema. At that moment, the intrinsic difficulty of the information was likely initially perceived as high.

For Trial 2, the concentration of pre-test scores was near 3 and, for Trial 3, the concentration of the pre-test scores was near 3 and 4; see Figure 5-f and Figure 5-i in Section 5.

I posit that the intrinsic difficulty of the incoming information imposed by the pre-tests for Trials 2 and 3 was initially perceived as low (easy to absorb or perform and to relate to existing knowledge about the topic) because students were accumulating knowledge continuously as they progressed. As a result, each of the pre-tests imposed a lesser intrinsic difficulty of the information and that difficulty of information is likely to be easier to process.

If the difficulty intrinsic to the information is perceived as lighter, less conscious cognitive effort would be required to process the information; therefore, the interventions themselves may not have been needed to help students process the information, which could explain the diminishing effect of the interventions. Also, this implies that in a real-life situation, these students who perceived the difficulty of the information as light might not receive the intervention because they might score above a threshold in the content quiz part of the environment. We can assume that there is much cognitive processing needed in Trial 1 but not so much in Trials 2 and 3. This seems to be related to desirable difficulty and teaching strategies, as will be discussed in a later section.

Future Research Recommendation: Change Topic for Each Trial

The value of this experiment is that it illustrates what happens when the information presented in a learning environment accumulates over time (Trials), with the result that the intrinsic difficulty of information imposed decreases as the learners' knowledge of the topic accumulates. This effect could not have been easily seen if the topics had been independent of each other. In that case, each topic would have imposed a high intrinsic difficulty each time students received new information about it (it is true that students would become increasingly familiar with the environment, but I speculate that this is unlikely to have more of an effect than the accumulation of topic knowledge). Future researchers should consider changing the topics for each trial to better observe the effect of interventions across different trials. For example, in Trial 1, students could learn about queuing theory, in Trial 2 students could learn about inventory management, and so on. That way, since the knowledge gained in each trial would not affect the students' understanding of the others, each trial would be more likely to be independent. I hypothesize that the value of presenting independent topics within the learning environment with interventions would be that the effects of the individual interventions themselves would be more clearly shown.

6.3.1.2 Effect of Repetition

Another possible explanation for the fact that once the students were in Trials 2 and 3, the intervention factors were no longer significant to learning gains is related to the effect of repetition, which, I posit, creates cumulative topic-related knowledge and familiarity with the environment. In the qualitative results, we see that 14 of the 24 students who commented specifically on repetition found it to be valuable for reinforcing their knowledge. This reinforcement could account for higher pre-test scores for these students, with the result that having an intervention would highly likely no longer be statistically shown to correlate with their learning gain. The demographic data supports this argument; many of the students who valued repetition had GPAs that tended to be relatively lower than those of the students who did not find it beneficial. If we refer back to Figure 5-p, the distribution of GPA for students who stated that the experience of the repetition was positive or negative, it seems reasonable to assume that students who found the repetition helpful might be those who needed extra repetition to get them used to the environment and material. For students who stated that the repetition was helpful, the average GPA was 3.30/4.00. For these students, I can assume that the new information was perceived as more difficult to relate to existing knowledge. That intrinsic difficulty of the information was likely initially perceived as high. The repetitions then enhanced their learning experience by reinforcing the knowledge and familiarity they needed to do better on the pre-test. Therefore, when new information was received in Trials 2 and 3, the intrinsic difficulty of information was less and the information was easier to process.

For the 10 students who stated that the repetition was not beneficial, the average GPA was relatively high (3.46/4.00); we can assume that students who reacted negatively were those who were simply able to understand the material faster and to infer further material more effectively. Therefore, it led to the perception that the repetition did not add value to their learning experience but instead just created annoyance, expressed as e.g., “Why am I doing this repeatedly?... I felt like it was unnecessary”. and “It almost felt too similar to the one before. It was just getting, like, frustrating”. It appears that for these students the repetition did not add value; however, the time they spent processing the repeated information could have added value by strengthening their knowledge about the topic. Thus, even though the students perceived the processing of similar information as affecting them negatively, the cumulative time they spent on the repeated material increased, which may have increased their familiarity with the topic.

The group of students who commented on repetition is a small sample of the students and thus it may not represent the whole population. However, it is reasonable to generalize; current literature states that repetition in education can produce a better learning experience (Bromage & Mayer, 1986). Additionally, Mayer and Johnson (2008) also conducted research exploring the redundancy effect in which they showed that the effect tends to reduce cognitive load processing. My results and conclusions partially agree with the published findings. However, my quantitative and qualitative findings suggest that repetition acts differently on different kinds of students.

Future Research Recommendation: Determine the Relationship between GPA and Students' Perception of Difficulty

Students with a higher GPA (i.e., greater than 3.5/4.00) are presumed to enter the trials with better knowledge and better ability to make connections and inferences, and thus they may require less repetition and the opposite may be true for students with lower GPAs. However, this research was not designed to test these assumptions about the relationship between response to repetition and GPA. Therefore, future researchers might attempt to establish a connection between GPA and perception of the intrinsic difficulty of the information, the ability to learn quickly, and the ability to draw inferences. This research may help determine how much repetition of materials and/or other content might be optimal for students.

6.3.2 Understanding How the Interventions Affect Student Learning

In this section, I focus on interpreting how the interventions may have affected student learning. I hypothesized that for my experiment: (1) the four intervention types (i.e., Audio-only, Text-only, Video, and Video+Text) would be key variables, (2) students' mean learning gains would not differ between different types of intervention (i.e., $H_0: \mu_{audio} = \mu_{text} = \mu_{video} = \mu_{mixed}$), and (3) students' learning experiences would differ among the interventions ($H_0: \mu_{audio} = \mu_{text} = \mu_{video} = \mu_{mixed}$). Obviously, there are more aspects of the students' experience than frustration; the level of frustration with the interventions, however, is the only quantifiable data that I collected.

What may not be obvious is that the second hypothesis depended on the validity of the first hypothesis. The multiple linear regression showed that this was partly true: the level of frustration with the intervention was significant. Since the first hypothesis was partly true, I just tested the second hypothesis using a Tukey test to determine whether the 2nd hypothesis were true or not. The discussion that follows will thus take a slightly different path from the usual.

My results support Hypothesis #1A to the extent that in the quantitative results, the level of student frustration with the intervention was one of the key variables affecting learning gains but the interventions were not in and of themselves significant. The quantitative results show that in Trial 1, the level of frustration with the four interventions was negatively significantly correlated to students' learning gains, as shown by a negative coefficient beta ($\beta = -0.225$). The fact that it is negative indicates that, at least by quantitative measurement, frustration is not providing value to students' learning gains. Also, the level represents the magnitude of external difficulty associated with the type of intervention: the higher the level, the more negative the learning gain scores. As explained above, since in Trial 2 and Trial 3 the intervention effects are diluted in the learning gain scores by the effects of repetition and inference, they were treated as negligible in the analysis. The interesting discrepancy between these quantitative data and the qualitative data will be discussed below.

My result also supports Hypothesis #1B; between different types of interventions, students' learning gains did not differ. As mentioned in the results section, Tukey's test reveals that there is no statistically significant difference between the learning gains associated with the interventions when each one is compared with each of the others; therefore, the results failed to reject our null hypothesis.

My quantitative result does not support Hypothesis #2; between different types of interventions, students' level of frustration with each intervention type did not differ. In Section 5.2.2, Tukey's test revealed that there is no statistically significant difference between the level of frustration associated with the interventions when each one is compared with each of the others; therefore, the results failed to reject our null hypothesis. This result indicates that the means of the level of frustration with the interventions are not different.

However, if we examine Figure 5-d, we can see the differences in learning experience associated with the four interventions, as reported in the Results chapter. To recap, Figure 5-d shows histograms of all four interventions separated by type. We can see that even though the means are approximately 2 for all interventions, the shapes of the tails differ slightly from type to type.

The survey questions about the level of frustration did not ask students to distinguish between the difficulty of the topic and the difficulty of the intervention medium. I am assuming that the level of the frustration is the product of the intrinsic difficulty of the topic + the external difficulty of the form that the information is coming in.

In the following section, I will discuss the qualitative and quantitative data relevant to each intervention and explain how students' learning experiences differ among the interventions.

I will also suggest a series of recommendations to take the best advantage of the nature of each intervention type.

6.3.2.1 Audio-Only Intervention: Low-Cost Option for Low Intrinsic Difficulty of the Information

On the basis of the results of this study, I suggest that the value of the Audio-only intervention could be high in situations when the intrinsic difficulty of the information is low, but the value is not high when the difficulty of the information is high.

In my study, I hypothesized that the Audio-only intervention would frustrate the students most among the interventions. The qualitative results reveal that most students who indicated that the audio was not helpful to them (35/53 students) said that the reason was that it was not visual enough. The quantitative data shows that Audio-only intervention is associated with a fairly high level of frustration. The reason for this could be that the topic in the experiment requires a large number of visuals and graphs, and students were frustrated that they did not have that. Furthermore, there could be an unspoken reason, i.e., that the information has to be mainly processed in the audio channel, which the qualitative data shows to be the less-favored channel among the participants.

It seems to me that it is unlikely that the Audio-only intervention will produce desirable difficulty when the source of the difficulty is the absence of visuals. Students' attempts to visualize the information by themselves might add so much extraneous difficulty that the result does not pay off well.

Research suggests that using the Audio-only intervention is not generally more beneficial than also combining audio with visuals (Mayer & Moreno, 2003), but if in designing an online learning environment we can see that the magnitude of the intrinsic difficulty of the information to be delivered is low, perhaps because the topic is very easy (an assumption that goes without saying) or is a fairly well-understood topic (such as a quick review), we can assume that audio delivery of information can be an excellent option because of the ease and low cost of production (Rossiter, Nortcliffe, Griffin, & Middleton, 2009). If, on the other hand, the material is new or very difficult (such that the intrinsic load would be heavy), I would recommend avoiding it as much as possible.

6.3.2.2 Text-only Intervention: Low-Cost and Reliable Option in Combination with Segmentation

On the basis of the results of this study, I suggest that the value of the Text-only intervention could be high in most situations if the designer of an online learning environment reduces the intrinsic difficulty of the information by appropriately segmenting the materials as needed.

I hypothesized that the Text-only intervention type would frustrate the students but not as severely as the Audio-only intervention. The qualitative results reveal that most students who indicated that the Text-only intervention was not helpful to them (27/52 students) said that the reason was that the quantity of text was perceived to be overwhelming. Also, the Text-only intervention exhibits a more even distribution of levels of frustration in 3, 4, and 5 than the Audio-only intervention. The overwhelming quantity of information presented all at once may have caused higher frustration (levels 3, 4, and 5).

However, the qualitative results also reveal that students who indicated that the Text-only intervention was helpful to them (9/52 students) said that the reason was that they were able to move through the material at their own pace. The fact that the Text-only intervention enabled them to process the information at their own pace may have caused less frustration (levels 1 and 2).

For students who were frustrated by the perceived overwhelming amount of information, the problem may be easily mitigated by further segmentation of the materials. One recent study found that chunking (segmenting) the video materials helped increase engagement and reduce extraneousness (Guo et al., 2014). It would be reasonable to assume that this would be true of other forms of learning materials. If text-based materials are segmented, then we would expect to observe a higher concentration of students' reports of their perceived levels of frustration in 1 and 2. However, in this case, the desirable difficulty that would be expected on the basis of time spent processing would probably not occur.

Some may argue that there may exist an optimal length that frustrates students just enough that the Text-only intervention can cause desirable difficulty. However, it is difficult to know what an optimal reading length is. According to Nanavati and Bias (2005), optimal length depends on each student's comprehension level, reading speed, method of movement (e.g., paging and scrolling), and eye movements. I would argue that unless technology can enable us to determine the exact length for each passage to achieve desirable difficulty through a wall of text, it would be preferable to use chunking for the time being and thus lessen the external difficulty.

In addition to the benefit identified by the students themselves, the Text-only intervention has other valuable features. It does not depend on the energy of the lecturer or the legibility of a presenter's handwriting. It can easily be edited and can accommodate the addition of visual aids, which are highly valued by students. Further, it is more reliably delivered than other types. According to a report on CBS on April 11, 2020, surveys undertaken in response to the upsurge in online learning sparked by the COVID-19 pandemic have revealed that many of US students have undergone bad online video learning experiences because of poor internet connections at their homes (CBS, 2020). In this kind of situation, given that we know that the text material has smaller file sizes than videos, we can see text serving as an excellent alternative option for online learning.

6.3.2.3 Video Intervention: Visual Features and Familiarity of the Form Minimize the Difficulty in Learning

On the basis of the results of this study, I suggest that the value of video intervention could be high irrespective of the magnitude of the intrinsic difficulty of content, but it requires careful consideration of many factors (e.g., cost of production, optimal video length, the energy of the lecturer, and availability of a reliable internet connection).

I hypothesized that the video intervention would frustrate the students the second least among the interventions. The qualitative results reveal that the students who indicated that the video was helpful to them (15/53 students) said that the reason was that it provided them with a visually supported detailed explanation that helped them to better understand the topic.

While indirectly, no cases were found in which students expressed frustration with video interventions, the quantitative data obtained from the 5-point Likert scale shows that video intervention leads to levels of frustration at level 3 or 4. The reason for this could have been that the video form was not frustrating even when the magnitude of intrinsic was high, because students may be very accustomed to learning via video. Also, in the interviews, they were not asked directly about their frustration but rather about their relative rankings of intervention types. The problems associated with text (e.g., too much text at once) and audio (e.g., not visual enough) may have been easier for them to isolate and identify. Thus, they might have been quicker to attribute the frustration to these features of the other forms than the video form. Contrariwise, problems with the primary lecture emerged (e.g., relatively unengaging delivery resulting in shortened attention spans) from the qualitative data, while the video intervention was delivered clearly and well. As discussed later in Section 6.5.3, features of the delivery could have affected the external difficulty but not have been associated with the intervention, which was the focus of the question the students were answering. As a result, it is likely that the difficulty of the topic was the source of the measured frustration.

If in designing an online learning environment, we can see that if the magnitude of the intrinsic difficulty of the information to be delivered is high, then the video may be a suitable option because the form itself seems to add little external difficulty. However, for video delivery of information to be excellent, designers have to ensure that the lecture is clear and concise, with well-designed visuals. Furthermore, for the video form to be effective at all, we have to assume that students have a strong connection to the internet.

6.3.2.4 Video+Text Intervention: Possibly Dependent on the Topic Difficulty

On the basis of the results of this study, I suggest that the value of the Video+Text intervention could be high irrespective of the magnitude of the intrinsic difficulty of content, but it requires careful consideration of features of both text and video forms, as described above. I hypothesized that the Video+Text intervention would frustrate the students the least among the interventions.

The results reveal that the large majority (48/53) of students who received the Video+Text intervention ranked it as their first or second choice, and a number of them specifically said that it helped them learn. The reason appears to be that it gave them a choice of intervention type and the option of using a different intervention if their first choice did not help them sufficiently. For these students, having an extra Text-only intervention helped them understand the video material better.

However, some students preferred the intervention with just a video to the Video+Text intervention. The reason appears to be that students felt that the text in the intervention was distracting them from learning. This is supported by the quantitative data, which shows that Video+Text intervention causes the frustration level to go up to level 3 or 4. Also, Video+Text has a higher probability of frustration in level 4 than the video alone. The higher level of frustration does not necessarily indicate that desirable difficulty would have been created, but the extra text might have exerted the same effect as repetition (see the effect of repetition above), which varies depending on the students.

For the Video+Text intervention, we can see that the Video+Text intervention is excellent if the magnitude of the intrinsic difficulty of the information to be delivered is high (difficult or relatively new), similar to the video. However, because some students appear to find the addition of text 'too much,' even when the difficulty of the main video is high, it is essential to make the text optional. Furthermore, the choice seems to be valued in and of itself by some learners, possibly because it gives them greater self-efficacy. It does seem clear that extra processing of text information can be mitigated by the choice of turning the text off for the students who do not value it. In this case, the extra text would not create a desirable difficulty because they could take the option to not use the material and still learn the material without using the text-based information.

However, on top of the demands of preparation of the video intervention, the Video+Text intervention requires an additional layer of effort. Not only does the Text-only intervention have to be good in and of itself, but it also has to work well in combination with the video. The two have to be seamlessly incorporated together.

6.3.2.5 Writing Intervention: Limitations, Desirable Difficulty Characteristics

On the basis of the results of this study, I suggest that the value of the writing intervention could be high because it forces engagement and can produce desirable difficulty. However, it has drawbacks that limit the situations in which it is likely to produce learning gains. One drawback is that students' misunderstanding of material can be reinforced because of a lack of structure and open-endedness of the writing task; and the second is that time spent on this activity can reduce short-term memory. Therefore, it is most effective when the perceived intrinsic difficulty of information is low, so that only well-understood information and concepts are reinforced by the performance of the task.

The quantitative results show that in Trial 1, the performance of the writing reflection was negatively significantly correlated to students' learning gains, as shown by a negative coefficient beta ($\beta = -0.235$). In the second and third trials, no statistical difference was found between the learning gains of those who performed the writing reflection and those who did not (as explained further in the aforementioned section).

There are three things that may have contributed to this scenario (negative learning gains, no learning gains).

- One is that time spent doing the task made them forget some details in the information and thus decreased the learning gain scores. According to Wickens, Gordon, and Liu (1998), the capacity of working memory is limited by how long information may remain and the fact that information will decay over time. Therefore, when the students were focused on doing the writing reflection, some key information about the topic may have been lost.
- A second is that the students may not have fully grasped the topic and this not fully understood or processed information could have gotten summarized in their own words.

This would result in the students' re-activating misunderstood information; according to Wickens et. al, information that needs to be remembered needs to be periodically reactivated (Wickens et al., 1998). Thus, the summarizing might cause the student to potentially deviate from learning the materials correctly.

- A third is that they were not familiar with the environment, as mentioned above.

It seems reasonable to posit that because of students' increasing familiarity and confidence as the trials proceeded, the effects of the effort put into the writing were diminished. That could account for the learning gain scores differences between Trial 1 and both Trials 2 and 3.

An additional explanation for statistically smaller learning gains for the group that had the writing intervention in Trial 1 may be that it limited those students to the use of only words to summarize a topic, "risk analysis," that depends heavily on the use of graphs, charts, etc., to illuminate the verbal explanations. As seen in the qualitative results, the majority of students have expressed the claim that visual aids help them and some also specifically expressed that they are visual learners. The concepts of risk analysis require extensive use of graphs and are easier to understand through visuals rather than through words. Therefore, words might not be enough for them to fully understand and explain the topic.

It is also possibly true that increased learning gains are not shown in the scores because, as reported by Dobson (2011), the desirable difficulty task may not produce immediate learning gains but long-term gains.

While quantitative results appear to suggest that the performance of the writing reflection did not add value to their learning, in the qualitative results, 30 of 35 students reported that the writing reflection was beneficial for a variety of reasons. This latter result supports one of the hypotheses of the study, that the writing reflection would improve self-reported memory.

However, the study results do not permit a simple conclusion that the writing intervention helps the learning overall.

On the basis of this study, I speculate that the writing reflection intervention has qualities of desirable difficulty. We can see that the writing intervention requires students to put in effort as they are summarizing and paraphrasing, which causes some difficulty for them. We can clearly see this ‘difficulty’ affecting them more severely in Trial 1, in that learning gains were statistically lower for the group who performed the writing reflection than for the group who did not. It appears that this writing reflection intervention becomes a ‘desirable’ difficulty in which expending effort may cause some unease at the beginning but gives the learners an overall good learning experience and an increase in self-reported memory. My conclusion agrees with the suggestion of Mueller and Oppenheimer (2014) that requiring students to synthesize and summarize content can serve as a desirable difficulty that leads to improved educational outcomes.

Recommendation:

I would recommend using the writing reflection intervention when the designers are confident that the students’ understanding of the topic is fully internalized in order for the intervention to have its full effect. In practical terms, this might necessitate not using writing intervention in the first portion of a learning experience but after students are familiar with the material. At the beginning of the online learning environment structure, less open-ended interactive tasks, such as content quizzes to engage students, can help them understand and reinforce what they are learning. After these interventions, I see a tremendous benefit in incorporating a writing reflection, but the task has to be clearly explained and the execution must not impose extraneous difficulty on students.

For complex topics that require visual aids, such as risk analysis or linear algebra, I recommend not only having a writing reflection but also including a drawing task to potentially improve students' learning and better suit their preferences. According to Hibbing and Rankin-Erickson (2003), evidence shows that with specific guidance as to what to draw, if students can create their own images, their understanding is increased. Because a number of students said that they liked the Video+Text intervention because of the variety of options, it might be beneficial to give students the option of choosing from different interventions (e.g., writing reflection, drawing task, or both) to help with their memory retention and improve their overall experience.

If an option were given to students to draw a single graph or figure, then they might be able to summarize and convey complex ideas of detailed information in a shorter amount of time. While drawing on the screen with a mouse might be harder for students than typing text on a computer, it could potentially add effort that might produce desirable difficulty.

6.4 Additional Factors That May Affect Students' Learning

In this section, I discuss the four additional factors extracted from the research findings that affected students' learning: inconsistency in terminology, interactive tasks, the energy of the lecturer, and visual aids. When I was designing the experiment, I was not looking to find these factors; therefore, I did not have hypotheses for them and did not ask explicitly about these factors in the interview questions. However, because these factors turned out to be significant in the student's students' learning experience, I discuss them and their effects here.

6.4.1 Inconsistency in Terminology

The results showed that inconsistency in terminology created confusion for students, but because they worked to figure out the meanings of the terms, they ended up understanding the terminology and content. When a student learning about a topic sees two different terms that seem to be used similarly, but that the student cannot be sure are synonymous, this may cause extreme confusion and difficulty in processing the information. The intrinsic difficulty of inconsistency of information requires a great deal of cognitive effort to be processed. The student may overcome the difficulty of perceiving the information, however, and if the student puts effort into figuring out two discrepant terms, it is likely to result in a long-term benefit. Also, if the student sees either of the terms in another place, it is less likely to create confusion, because the he or she will have incorporated the new knowledge.

If the student does not overcome the difficulty of perceiving the discrepant information, then there is a good chance that information that is being delivered will remain not understood.

Research has shown that inconsistency in terminology can produce extreme extraneousness. Grünewald, Meinel, Totschnig, and Willems (2013) reported a finding on inconsistent definitions and found that when students encountered contradictory definitions, they were dissatisfied. It is possible that for students in my study, the discrepancy between terms or definitions was too strong and students could not overcome the difficulty, causing the information to remain not understood. My research shows that it does produce difficulties for students but also that the degree of difficulty varies from one student to another. Some students may be able to handle inconsistency in terminology and figure out the connection, but for students unable to manage it, it will be a highly external difficulty. On the basis of this, my recommendation is that when designing the online learning environment, a designer should try to avoid inconsistent terminology as much as possible because there is a high risk that it will cause enough extreme difficulty that the information may not be processed and understood despite students' efforts. In other words, desirable difficulty may not be produced, and no value is added. Furthermore, if the inconsistency in terminology is frequent in an existing set of learning modules, learning aids (interventions) may be needed. I recommend text because of the ease in changing the content in the text materials, as opposed to changing materials in audio and video learning materials, which requires more post-production edits.

6.4.2 Interactive Tasks

It is possible that the interactive tasks in my experiment were beneficial because each question or task revealed a little gap in the students' knowledge and required them to put in their cognitive work to fill the gap. In the qualitative results, students said that participation in interactive tasks (writing reflection and content quizzes) positively influenced their learning and made them more confident about their own learning experience. Performing these tasks did not create a great deal of external difficulty, but it did require enough effort to overcome it and make connections between incoming information and existing knowledge to have possibly resulted in desirable difficulty. For example, when students were asked to do the writing reflection, part of the intrinsic information of the topic may already have been understood such that the topic was superficially grasped but not fully internalized. Then, when students became engaged with the task, having to use their own existing vocabulary to summarize and capture new information may have helped them internalize the intrinsic information more and strengthened the connection of the information to their internal knowledge schema.

Also, the content quizzes, which contained real-life example questions, required students to put in cognitive work to extend what they had learned to a new situation to fill the gap in knowledge. Overcoming the external difficulty of this seems to help improve learning. Support for this argument can be found in D. Zhang et al. (2006), where it was reported that in an online learning environment where students were asked to interact with the material by annotating important notes on a video, this interaction was shown to improve students' learning. We can posit that such interactive tasks made students overcome a small amount of external difficulty at short intervals to constantly keep internalizing the intrinsic information. This information was manageable because the intrinsic information was segmented into smaller chunks and they had to overcome smaller amounts of external difficulty at a time. Both my results and the literature indicate that interactive tasks improve the learning experience by incrementally building students' confidence in their learning. Also, both indicate that the benefit results from the cognitive efforts that the students must exert to complete these activities. This result also sheds light on the underlying processes that account for various cognitive efforts and shows why constantly revealing gaps and getting students to fill them can produce better learning.

6.4.3 Energy of the Lecturer

In the qualitative results, students expressed that their perceptions of the energy and engagement of the lecturer have a significant influence on how well and how long they can pay attention. Students' perception of the energy level of the lecturer affects their perception of whether or not videos are too long and also their ability to maintain their attention for as long as they need to. For example, students stated that the low energy of the primary lecturer in the experiment made it difficult for them to continue paying attention.

The role of energy and enthusiasm is highlighted in the audio or video learning materials, where the teachers' enthusiasm and vocal qualities are significant features. I posit that the less energized and engaged the lecturer is, the greater the perceived external difficulty; therefore, the higher the likelihood that the intrinsic information will be processed less easily. If this is the case, then there is a higher likelihood that more cognitive effort will be required to process the information.

In contrast, the more engaged the lecturer is in presenting, the less cognitive effort is required for students to process the combined intrinsic and external difficulty of information. Some support for this argument can be found in (Guo et al., 2014), where it is reported that the students' engagement increases when the lecturer's speaking rates increase. While the rate of speech does not necessarily mean higher energy, typically faster speech is associated with higher energy. In this case, the higher energy would lead to higher attention, which leads to a higher likelihood that students can focus their cognitive efforts on taking in the intrinsic information because they do not have to overcome as much external difficulty. My result agrees with their findings in reverse: low-energy speech [delivery] was associated with poor attention and diminished engagement.

I would argue that it would be difficult for this factor to encourage students to put more effort into learning the material. In other words, it will be less likely that the factor will produce desirable difficulty and could ultimately cause students to simply exit the learning environment. This factor would increase the probability that students will perceive too much external difficulty in their learning and feel that they cannot overcome it. This is just speculating on the basis of the qualitative results, but these show that the negative effect of lower energy and engagement is severe for students; it causes tremendous external difficulty that makes them lose focus and lose attention. If that occurs, the intrinsic information will fail to be processed. Therefore, because the likelihood of desirable difficulty is not high, I would suggest that it would be preferable to ensure that the lecturer is energetic and engaging.

6.4.4 Visual Aids

Overall, visual aids appear to provide a better learning experience for the majority of students and are important for ensuring that students get an early grasp of the topic. In the qualitative results, many students expressed that more visuals (graphs, pictures, and videos) are helpful in their learning, and some directly stated that they self-identify as visual learners. It is possible that students perceive visual aids as valuable tools in their learning experience because they simply have a preference for presentations in which the audio or text delivery is supported by visual aids.

From the data, we can posit that the more visual aids in the lecture, the higher the probability that external difficulty will be reduced. As shown in Figure 5-d in section 5, we can clearly see the evidence that, with the same intrinsic information, we can observe that the distribution of frustration for the Video intervention has higher concentration in levels 1 and 2 and the distribution of frustration for the Audio-only intervention has wider variation. There was no difference in the mean statistically; however, the Audio-only and the Video levels of frustration support the claim that more visuals result in reduced variability in external difficulty.

There is an assumption here that goes without saying, i.e., that relevant and appropriate amounts of visual aids would reduce variability, while irrelevant and excessive visual aids could increase the variability. From the designer's perspective, the video form may not be fully optimal, considering the time, money, and effort spent in producing learning materials in this form. Furthermore, their effectiveness depends a great deal on the clarity, energy, and engagement of the lecturer, as discussed in the previous section. Mayer and Moreno (2003) suggest reducing the cognitive load to optimize learning capacity; to do that, one should use appropriate levels of both visual and auditory channels. My research corroborates existing research by demonstrating that for engineering students, the designer of an online learning environment can expect better results with appropriate usage of visual aids.

6.5 Summary of Insights and Takeaways

There are many takeaways from this study. One is a description of how different interventions affect the learning experience. Another is the identification of a number of factors that interact with the interventions and affect the learning. This study merges quantitative and qualitative information to reveal what really influences students' learning in an online learning environment and does show, in accordance with the literature and the findings of the pilot study, that frustration is an important piece in students’ learning. This study also presents practical considerations and recommendations based on multiple perspectives (e.g., students’, teachers’, and online learning designers’).

Chapter 7 Conclusion

The specific problem addressed in this study was the lack of a research-based understanding of how interventions function in the online learning experience. The purpose of this study is to: (1) provide an in-depth exploration of factors that affect students' learning experience in an online learning environment; (2) illuminate the features of interventions that affect undergraduate engineering students' online learning experience; and (3) investigate relationships between factors in order to demonstrate mutual influences, both positive and negative. As the literature suggests, online learning is playing an increasingly important role in broadening access to high-quality education throughout the world, and more research on helping to design online materials is needed to help students learn better.

For this study, I first designed an online learning environment with interventions based on recommendations derived from cognitive load theory. The experiment was designed to simulate the online learning environment with interventions. I recruited 70 undergraduate engineering student participants who had no knowledge about the topic to be presented. During and after the participants' completion of the experimental procedure, I collected qualitative and quantitative data about their experiences. To analyze the data, quantitative analysis (multivariate regression analysis, descriptive statistics, Tukey's test) and qualitative analysis (thematic analysis) were performed and merged to answer the following research questions:

- Research question 1: What factors in online learning environments affect learning gains (i.e., measured difference between the post-test and the pre-test scores) for undergraduate engineering students?

- Research question 2: What factors in online learning environments affect the learning experience for undergraduate engineering students, and, specifically, what factors produce desirable difficulty?
- Research question 3: What factors in online learning affect undergraduate engineering students' self-reported memory?

7.1 Summary of Results: An Overarching Theme

In this study, aspects of the learning process that are peculiar to the online learning environment were examined, but a major key finding that emerged about online learning is that learning, regardless of the form in which it occurs, follows the same fundamental processes as it has for thousands of years: in order for learners to grasp new information, they need to overcome some frustration. However, if students are going to overcome frustration, the frustration must be manageable and the delivery of the information must be suited to their needs.

In order to bring that about, successful educators must be able to understand where the students are coming from (e.g., their knowledge state, socioeconomic background, demographic background) and provide them with an appropriate level of frustration in learning tasks. This is particularly important in online learning because it is typically self-directed. Also, this work points out a strength of online learning environments: they are poised to take advantage of new developments in artificial intelligence (machine learning). Since machines will be able to continuously collect data from students, with more development, they will become increasingly better at analyzing students' needs, and be better able to select and provide appropriate materials adaptively for individual students in online learning systems.

The main contributions of this research are: (1) identification of some important considerations for designers who are using interventions, (2) identification of some factors that play a large role in students' learning that might have not been immediately obvious, and (3) a set of practical recommendations for designers of online learning environments. The detailed findings for each of the research questions are summarized briefly below:

7.1.1 Summary of Results for Research Question #1

For Trial 1, the level of frustration with the four types of intervention (Audio-only, Text-only, Video, Video+Text), performance of the writing reflection, and pre-test scores affected learning gains. For Trials 2 and 3, only pre-test scores were significant for learning gains; the frustration level with four types of interventions and the performance of the writing reflection were not significant. As expected, the quantitative data showed that the forms in which the students received the new information (i.e., Audio, Text, Video, Video+Text) affected students' learning gains, but indirectly. Also as expected, in the quantitative results, the performance of the writing reflection was significant, but unexpectedly, it was negatively correlated with the learning gains.

7.1.2 Summary of Results for Research Question #2

Each intervention (the four types of intervention and the performance of a writing reflection) affected students' learning experiences in different ways. Surprisingly, in light of the negative correlation reported above, many students expressed in the qualitative (interview) data that the performance of the writing reflection was helpful in their learning. Additionally, analysis of the qualitative data in light of the quantitative results revealed unexpectedly that other factors that had not been intuitive for me to measure at the time of the experiment also affected the learning experiences significantly (e.g., the instructor's energy and vocal quality, inconsistency in terminology, presence of visual aids, repetition).

In regard to desirable difficulty, as expected, there is evidence that the Text-only intervention did produce desirable difficulty. However, the Audio-only intervention apparently did not. The writing reflection intervention also produced some desirable difficulty, which might account for the unexpected negative correlation with learning gains (as a result of the spacing effect).

Repetition and inconsistency of terminology can be conjectured to have contributed to frustration in some instances but did produce desirable difficulty in some other cases. Repetition was shown to have had both positive and negative effects on students' learning experience as expressed qualitatively; the difference appeared to be related to students' presumed ability to grasp the material quickly. The presumption seems to relate to the students' GPAs (collected in the descriptive statistics); on the whole, students whose reported GPAs were slightly higher tended to find repetition unhelpful and frustrating while, on the whole, students whose reported GPAs were slightly lower said that it was helpful. This finding suggests a possible correlation that might be further investigated.

7.1.3 Summary of Results for Research Question #3

While I expected that some interventions would affect self-reported memory, unexpectedly, results suggested instead that interactions, the energy of the teacher, and self-efficacy of visual learners increase self-reported memory. These three factors play key roles in reducing unnecessary (non-value added) frustration and helping students learn most effectively with respect to long-term learning gains.

.

7.2 Implications of This Work and Proposed Recommendations

The findings of this study have numerous implications for the design of online learning environments for educators and instructional program designers. Broadly, these are related to the key concepts of desirable difficulty and management of the cognitive load.

First, on the basis of an understanding of desirable difficulty in OLE, educators can design learning tasks to be more or less difficult to accommodate students' learning preferences. In order to effectively use the concept of desirable difficulty, (i.e., apply it to interventions or even in general learning), we have seen that it is better to challenge the students (i.e., give them a difficult task as an intervention) only when they have a solid understanding of the topic (indicated by performance on a test), especially in self-directed learning. Students may experience frustration that is too great for them to overcome when they are faced with difficulty before they are adequately prepared for it, especially if (as in self-directed learning) they have nowhere to turn for help with the material.

Second, online program designers can help manage the cognitive load by ensuring that learners' prerequisite knowledge is adequate. Given the key assumption of cognitive load theory, that working memory is limited, giving learners difficult tasks while they are dealing with an intrinsically difficult topic that they do not yet understand will impose a burden because the learners must focus some of their cognitive capacity on addressing a difficulty that is not likely to be manageable at that point. Therefore, it is crucial for designers to ensure that learners have basic knowledge obtained first through appropriate materials; then they can focus on the new information that it is given to them.

7.2.1 Recommendations for Practice

On the basis of the implications, the following practices are recommended:

1. For high-intrinsic load materials (a difficult topic or very new topic), the designer should use visual learning materials (video form and visual aids) to convey the information to learners, since this makes optimal use of the channel that undergraduate engineering students typically favor. Even though videos may take more time and/or cost more to make, the benefit for students may compensate for those costs.

2. For low-intrinsic load materials (review of already presented materials or summaries), Audio-only interventions may be appropriate because they are easier to make and almost certainly less costly.

3. To accommodate students who have a poor/unreliable internet connection due to various circumstances, a Text-only intervention is an acceptable option, as it provides good information and permits self-paced reading while requiring relatively little bandwidth. However, it is effective only if the materials are broken down into manageable segments.

4. For video interventions, designers should offer additional text information to get the most out of its value. Also, designers should include a turn-on/off option for the text information to avoid information overload and undesired mixed modality.

5. Especially in an online learning setting where the learning is mostly self-directed, writing reflection interventions should be used only when students have acquired a solid understanding of the topic already to avoid the possibility of reinforcing incorrect information.

6. Designers should follow traditional instructional recommendations on providing information clearly, displaying high energy, using clear visuals, and including interactive tasks. I would recommend not designing desirable difficulties that are in conflict with these practices (for example, using hard-to-read fonts, having unnecessary animations, etc.).

7. Designers should carefully consider the effects of repetition and inconsistency in terminology depending on circumstances. Repetition may be ideal for someone who does not have adequate knowledge, but it may frustrate students if they already understand the information. Also, inconsistency in terminology is a high risk and high reward concept where if students can overcome their frustration and reconcile inconsistencies, it may benefit them in the long term; but if they are unable to do so, it can potentially create long-term disadvantages that propagate throughout their learning.

7.3 Limitations

There were several limitations to this study design that imposed constraints on the conclusions I could draw from the data:

1. This experiment did not change topics from one trial to the next to illustrate what happens when the information presented in a learning environment accumulates over time (trials), with the result that the difficulty imposed by the incoming information decreases as the learners' knowledge of the topic accumulates. Changing the topic between trials would have enabled me to get better information about the effect of the interventions.

2. The majority of the student subjects were industrial and mechanical engineering undergraduate students. Using a broader spectrum of engineering majors would have allowed me to extrapolate a better understanding of engineering learning.

3. I formulated one of the hypotheses for Research Question #1 based on the assumption that the first hypothesis was true prior to confirming it. The effect of that was that it could have taken the study in the wrong direction; however, luckily, it helped me understand better the effects of interventions and how the use of interventions is related to the learning experience. However, for future work, I will make sure that I do not jump to a conclusion and that I make my hypotheses based on a research question asked.

4. This research did not use a Likert scale to measure perceptions related to the writing intervention and other factors because some were unexpected factors; I did not expect the writing task to have much of an impact on students' frustration in learning. For future work, I would suggest using a Likert scale to measure students' responses. This would likely enable better comparisons between factors and allow us to see the differences in the effects.

5. The experiment measurements did not isolate different degrees and sources of frustration and there were unrepresented perception measurements, particularly "0 = no frustration". This also resulted from the unexpectedly significant role of frustration in the experiment. Also, I was not able to prepare appropriate interview questions and measurement tools to separate frustration due to intrinsic difficulty of the information from that due to external difficulties. To avoid these problems in the future, I plan to carefully prepare Likert scales that represent all levels of perception (frustration) and also restructure the interview questions to elicit the different causes of frustration.

6. Because of time constraints, an experiment with a control group who did not have interventions was not conducted. Having a control group would have enabled the experiment to measure learning gains that are presumably exclusively due to interventions.

7.4 Recommendations for Future Work

To expand on the work of this study, a first piece would be to investigate whether there is a correlation between perception of intrinsic difficulty of information and learner demographics (e.g., GPA) so that these data could be an input to an online learning system to provide students with appropriate materials. With these additional capabilities, my model could provide better and more appropriate materials/interventions for a student in accordance with his/her needs or goals in online learning. These would not be the only determining factors for perception of intrinsic difficulty of information, so bias would not be a concern.

Secondly, as mentioned in the Methods, it should be noted that the possible effects having to do with pure language concerns, e.g., effects of English as a second language, on the part of either the lecturers in the learning materials or the participants, have been treated as negligible for this study. In recognition of this possibility, I reported citizenship status in the Descriptive Statistics section because of the higher likelihood that participants who were non-citizens might have second language concerns, but I have not explored that in depth in this research. I intend to incorporate this consideration in subsequent studies.

If I were to refine this experiment in light of the limitations noted above, I would make the following modifications:

- I would like to run this experiment again in a real adaptive learning environment, rather than a simulated one, with many more demographically varying subjects.

- I would also like to test different fields other than engineering with a different subject group to see if it produces the same results. Also, I would like to test different topics for each trial; this would ensure non-cumulative knowledge and provide better insights about the effects of interventions.
- As mentioned above, I would like to have an additional test for this research with a control group who did not have an intervention. This would allow us to see more clearly whether the learning gains were happening because of the interventions or not.
- In conclusion, all of these suggestions might be expected to provide researchers in the field a better data-based understanding of the cognitive load involved in learning in a defined body of material and ways to identify difficulty that helps in the learning process, which I believe will contribute to the future development of more effective ways to help students learn.

Appendices

Appendix A Screenshots of Sample Interventions

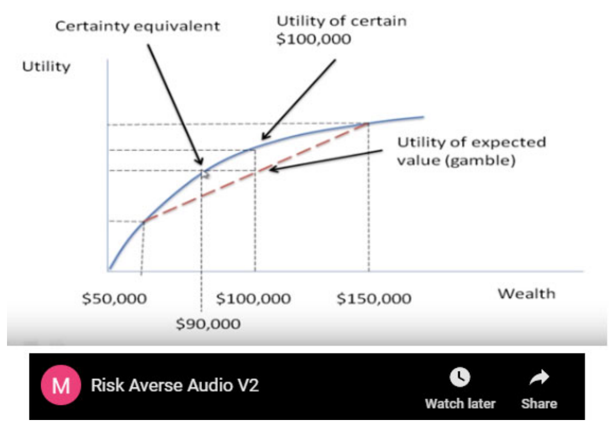
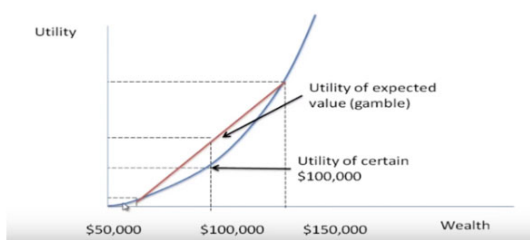


Figure A-1: Audio Intervention Sample



They exhibit this increasing marginal utility of wealth so the curve is bowed upward.

As shown in the Figure, the $100,000 with certainty actually gives them less utility then the expected $100,000 they get from taking the gamble. This is a kind of person who wants to take the gamble than a certain amount. There are examples of this, usually lower levels of wealth:

- people who play the lottery
- people who go to the horse to the racetrack and bet on horses
- people who rather go to Las Vegas and take the 50% chance that I lose $100 and 50% chance that I win $100 than to take the certainty of just doing nothing.

These are people who would rather take the gamble than to take the certain amount. We see this example in certain kinds of behavior but usually smaller levels of wealth. The utility function might be current bowed up at low levels of wealth and then showing diminishing marginal utility of wealth at higher levels of wealth. This gives you an idea of why some people take a gamble and why some people don't and in fact if we sort of scroll back here to the risk-averse individual

Figure A-2: Text Intervention Sample

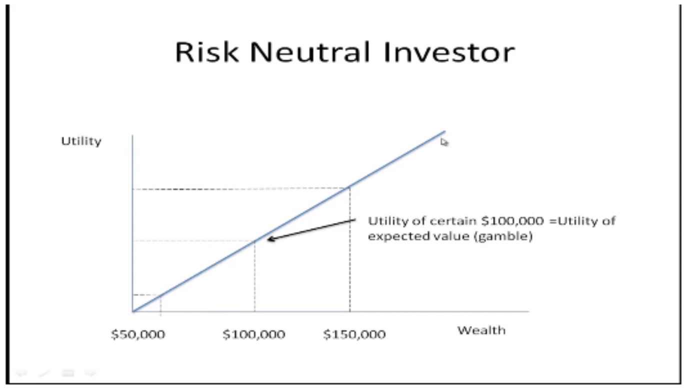
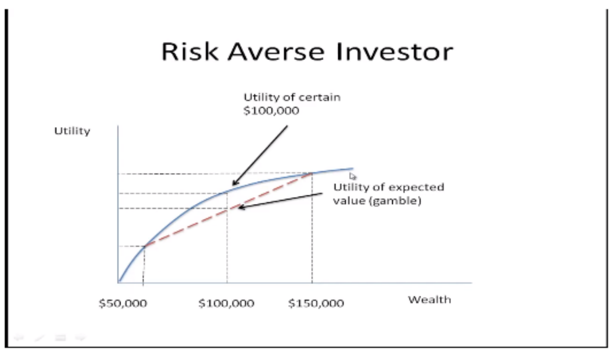


Figure A-3: Video Intervention Sample



Risk averse investors don't want to take the gamble. They're not willing to take the risk **The blue line** represents the risk averse investor's utility function and at each level of wealth, they have a certain level of utility or satisfaction.

For example, you can see that going from a 100k to 150k increased utility by a smaller amount. But going from 50k to a 100k increased utility by a large amount. That's the **decreasing marginal utility of wealth.** That is, the first $50,000 increase or the increase from 50k to 100k gave them a greater jump in satisfaction than going from a 100k to 150k.

The **red dashed line** represents the utility you get from gambling. In this case, because the expected value of the gamble is a $100k, we can draw up to this dashed line and draw across this would be their level of utility. Utility of expected value is lower than the 100k with certainty. In other words, this risk adverse investor would rather have a 100k with certainty than taking the gamble of 50% chance of getting 50k or 50% chance of getting 150k. Therefore, the utility from the certainty of a 100k is greater than the expected utility

Figure A-4: Video+Text Intervention Sample

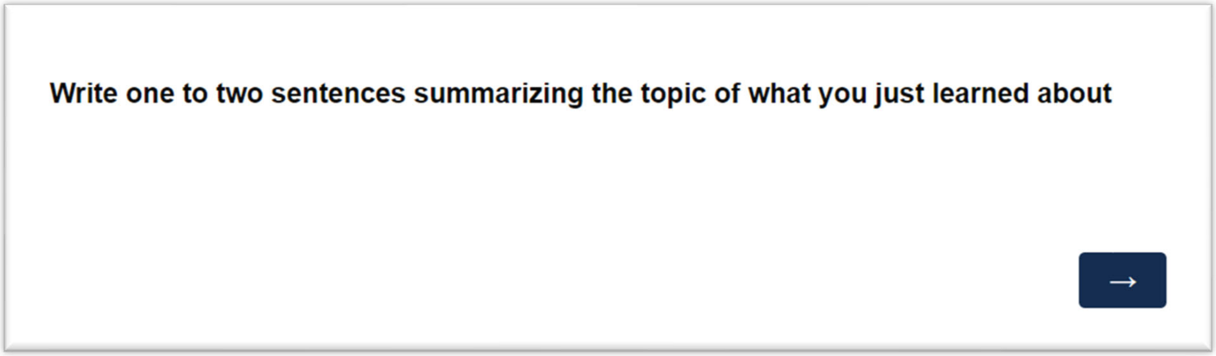


Figure A-5: Writing Intervention Sample

Appendix B Recruiting Email

Hello,

My name is Seok-Joo Kwak and I am an Industrial and Operations Engineering PhD student. I am seeking participants for a research study on adaptive learning environment (UM IRB approval HUM00161103). Your participation in this study will consist of completing Decision Analysis modules at a designated campus location. The results of the experiment will be used to learn about factors to improve the adaptive learning environment.

Participation: To participate in this study, you must be at least 18 years or older AND a University of Michigan Engineering student.

Compensation: Participants will receive a $20 MasterCard gift card for participation in this study.

Duration: Approximately 1 hour.

Location: 1110 ERB (2200 BONISTEEL BLVD Ann Arbor, MI 48109)

Data Collection Period: April 10, 2019 - July 30, 2019

If interested, please send an email to seokjook@umich.edu. You will be then contacted to confirm your eligibility and to schedule a time to participate.

Thank you!

Appendix C Demographics

Table A-1: Experiment Demographics Sex

Sex	Count
Female	29
Male	41
Grand Total	70

Table A-2: Experiment Demographics Major

Major	Count
Mechanical Engineering (ME)	36
Industrial and Operations Engineering (IOE)	29
Biomedical Engineering (BME)	3
Nuclear Engineering and Radiological Sciences (NERS)	1
Materials Science and Engineering (MSE)	1
Grand Total	70

Table A-3: Experiment Demographics Citizenship Status

Status	Count
United States citizen	63
Neither a United States citizen nor a permanent resident	5
United States permanent resident	2
Grand Total	70

Table A-4: Experiment Demographics Race

Race	Count
White/Non-Hispanic	44
Asian	17
Hispanic or Latino	5
Black or African American	4

Appendix D Interview Protocol

Interview Protocol

Interview after module completion:

1. What did you think about the overall process of completing the module?
2. What were some pros and cons about the modules you just watched?
3. How would you rank interventions (worst to best)? and Why?

I want you to think back to one of the best lectures you have had in your academic experience. Now thinking about that, please describe:

4. What did you like about that lecture experience?
5. What did you dislike about that lecture experience?

Now, think about your experience in completing the module.

6. What did you like about your experience completing the module?
7. What did you dislike about your experience completing the module?
8. How would you compare your lecture experience to your module experience?

Describe a time in this process in which you felt frustrated doing the modules.

   a. What was frustrating about it?
   b. Please tell me a little bit more about why that was frustrating.

9. Describe any moments during the modules when the video failed to explain concepts thoroughly.
10. Describe any moments where you could not pay attention.
    a. What was hard to pay attention to?
    b. Why do you think it was hard to pay attention at this time?
11. Please feel free to share any thoughts you had about this overall process.
12. Is there anything else that you would like to add?

Thank you for participating in this study. For your participation, you will receive a $20 MasterCard gift card. Please fill out this form.

Appendix E IRB Exempt Information Sheet

INFORMATION SHEET
TITLE OF THE RESEARCH PROJECT
HUM00161103

Principal Investigator: **SeokJoo Kwak, PhD Candidate, University of Michigan**
Faculty Advisor: **Joi Mondisa, PhD Assistant Professor Department of Industrial and Operations Engineering**

You are invited to participate in a research study about examining factors and designs that improve undergraduate engineering student learning and experience in an adaptive learning environment.

If you agree to be part of the research study, you will be asked to watch 3 Topics in an Operations Research Course. You will perform the following steps:

1. Take a pre-test about a Topic.
2. Watch a video on the Topic
3. Approximately 3-9 minutes into the video, the system displays a short multiple-choice quiz about the topic.
4. You will be provided with various types of treatments module.
5. Take a post-test about Topic
6. Take survey questions in Qualtrics about your experience

After you are done with the Topic 1 section, you will go through the whole cycle with Topic 2 and 3 next. You will take demographic survey followed by an interview.   Entire process will you take around 1 hour in the lab.

Benefits of the research: Although you may not directly benefit from being in this study, others may benefit because the empirical findings of this research may result in an increased public understanding of human in adaptive learning in the massive open online courses learning.

Risks and discomforts: There are no direct risks to the public or community, which could result from this research. Voice recordings will be transcribed into word documents, so they are not identifiable. Voice recordings will be deleted immediately after the interview.

Compensation: You will be given a $20 gift card to thank you for your participation in the study.

Participating in this study is completely voluntary. Even if you decide to participate now, you may change your mind and stop at any time. You do not have to answer a question you do not want to answer. If you decide to withdraw before this study is completed, your information will be destroyed.

I/We will protect the confidentiality of your research records by changing names to pseudonyms and voice recordings will be transcribed into word documents, so they are identifiable. Voice recordings will be deleted immediately after the interview. Furthermore, transcribed word documents will be stored in Seok-Joo Kwak's locked storage. During the experiment, you will enter a soundproof lab room with only a laptop computer with a video module open. PI, SeokJoo Kwak, will be in the room only to assist the participant if there is a troubleshoot and won't be able to see the screen.

For a detailed description of CoC protections and exceptions to those protections, please refer to Attachment A at the end of this document.

Information collected in this project may be shared with other researchers, but we will not share any information that could identify you.

If you have questions about this research study, please contact **SeokJoo Kwak (810) 333-2157 or Joi Mondisa (734) 647-8720**

The University of Michigan Institutional Review Board Health Sciences and Behavioral Sciences has determined that this study is exempt from IRB oversight.

References

Abbakumov, D., Desmet, P., & Van den Noortgate, W. (2020). Rasch model extensions for enhanced formative assessments in MOOCs. *Applied Measurement in Education, 33*(2), 113-123.

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive science, 26*(2), 147-179.

Allen, I. E., & Seaman, J. (2017). Digital Compass Learning: Distance Education Enrollment Report 2017. *Babson survey research group*.

Alsanad, S. M., Howard, R. L., & Williamson, E. M. (2016). An assessment of the impact of herb-drug combinations used by cancer patients. *BMC complementary and alternative medicine, 16*(1), 393.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes.

Bárcena, E., Read, T., Martín-Monje, E., & Castrillo, M. D. (2014). Analysing student participation in Foreign Language MOOCs: a case study. *EMOOCs 2014: European MOOCs Stakeholders Summit*, 11-17.

Baturay, M. H. (2015). An overview of the world of MOOCs. *Procedia-Social and Behavioral Sciences, 174*(12), 427-433.

Beck, J. E., & Woolf, B. P. (2000). *High-level student modeling with machine learning.* Paper presented at the International Conference on Intelligent Tutoring Systems.

Belanger, Y., & Thornton, J. (2013). Bioelectricity: A quantitative approach Duke University's first MOOC.

Bjork, R. A. (1994). Institutional impediments to effective training. *Learning, remembering, believing: Enhancing human performance*, 295-306.

Bosch, N., & D'Mello, S. (2017). The affective experience of novice computer programmers. *International Journal of Artificial Intelligence in Education, 27*(1), 181-206.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*: sage.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education, 15*(4), es6.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.

Bromage, B. K., & Mayer, R. E. (1986). Quantitative and qualitative effects of repetition on learning from technical text. *Journal of educational psychology, 78*(4), 271.

Brown, S. J. (2015). Personalized online learning management system and method: Google Patents.

Capdeferro, N., & Romero, M. (2012). Are online learners frustrated with collaborative learning experiences? *International Review of Research in Open and Distributed Learning, 13*(2), 26-44.

Casey, D. M. (2008). The historical development of distance education through technology. *TechTrends, 52*(2), 45-51.
Castillo, N. M., Lee, J., Zahra, F. T., & Wagner, D. A. (2015). MOOCS for development: Trends, challenges, and opportunities. *International Technologies & International Development, 11*(2), 35.
Cavanaugh, C., & Clark, T. (2007). The landscape of K-12 online learning. *What works in K-12 online learning. Eugene, OR: International Society for Technology in Education*.
CBS (Producer). (2020, June 1 2020). Nationwide school closures make learning difficult for students. Retrieved from https://www.youtube.com/watch?v=iMbEmsb_UHE
Centra, J. (2009). TM Differences in Responses to the Student Instructional Report: Is It Bias? Retrieved from http://www.ets.org/Media/Products/SIR_II/pdf/11466_SIR_II_ResearchReport2.pdf
Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Undesirable difficulty effects in the learning of high-element interactivity materials. *Frontiers in psychology, 9*, 1483.
Chen, O., Woolcott, G., & Sweller, J. (2017). Using cognitive load theory to structure computer-based learning including MOOCs. *Journal of Computer Assisted Learning, 33*(4), 293-305.
Chen, Y., Culpepper, S. A., Wang, S., & Douglas, J. (2018). A hidden Markov model for learning trajectories in cognitive diagnosis with application to spatial rotation skills. *Applied psychological measurement, 42*(1), 5-23.
Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. (2013). The MOOC phenomenon: who takes massive open online courses and why? *Available at SSRN 2350964*.
Conrad, D. L. (2002). Engagement, excitement, anxiety, and fear: Learners' experiences of starting an online course. *The American journal of distance education, 16*(4), 205-226.
Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW: Citeseer.
Cowan, J. (2014). Noteworthy matters for attention in reflective journal writing. *Active Learning in Higher Education, 15*(1), 53-64.
Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative*: Prentice Hall Upper Saddle River, NJ.
Cutrell, E., O'Neill, J., Bala, S., Nitish, B., Cross, A., Gupta, N., . . . Thies, W. (2015). *Blended learning in Indian colleges with massively empowered classroom.* Paper presented at the Proceedings of the Second (2015) ACM Conference on Learning@ Scale.
De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional science, 38*(2), 105-134.
De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113-140.
Despujol, I. M., Turró, C., Busquéis, J., & Cañero, A. (2014). *Analysis of demographics and results of student's opinion survey of a large scale mooc deployment for the spanish speaking community.* Paper presented at the 2014 IEEE Frontiers in Education Conference (FIE) Proceedings.
DeWall, C. N., Baumeister, R. F., & Vohs, K. D. (2008). Satiated with belongingness? Effects of acceptance, rejection, and task framing on self-regulatory performance. *Journal of personality and social psychology, 95*(6), 1367.

Dewey, J. (1933). How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process Vol. 8.
DiGangi, C. (2017). The average student loan debt in every state. Retrieved from https://www.usatoday.com/story/money/personalfinance/2017/04/28/average-student-loan-debt-every-state/100893668/
Dobson, J. L. (2011). Effect of selected "desirable difficulty" learning strategies on the retention of physiology information. *Advances in Physiology Education, 35*(4), 378-383.
Downes, S. (2008). Places to go: Connectivism & connective knowledge. *Innovate: Journal of Online Education, 5*(1), 6.
Freedman, J. (2013). Why American Colleges Are Becoming a Force for Inequality. Retrieved from https://www.theatlantic.com/business/archive/2013/05/why-american-colleges-are-becoming-a-force-for-inequality/275923/
Gaebel, M. (2013). MOOCs Massive open online courses. *EUA Occasional papers*.
Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction, 28*(2), 127-203.
Grabmeier, J. (2015). 70 percent of college students stressed about finances. Retrieved from https://news.osu.edu/news/2015/07/01/financial-wellness/
Gravetter, F. J., & Forzano, L.-A. B. (2018). *Research methods for the behavioral sciences*: Cengage Learning.
Grünewald, F., Meinel, C., Totschnig, M., & Willems, C. (2013). *Designing MOOCs for the support of multiple learning styles.* Paper presented at the European conference on technology enhanced learning.
Guo, P. J., Kim, J., & Rubin, R. (2014). *How video production affects student engagement: an empirical study of MOOC videos.* Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.
Gutiérrez-Rojas, I., Alario-Hoyos, C., Pérez-Sanagustín, M., Leony, D., & Delgado-Kloos, C. (2014). Scaffolding self-learning in MOOCs. *Proceedings of the European MOOC Stakeholder Summit, 2014*, 43-49.
Harrington, N. (2005). The frustration discomfort scale: Development and psychometric properties. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 12*(5), 374-387.
Hatzilygeroudis, I., & Prentzas, J. (2009). A survey on recent patents regarding intelligent educational systems. *Recent Patents on Computer Science, 2*(3), 214-222.
Heitmann, S., Grund, A., Berthold, K., Fries, S., & Roelle, J. (2018). Testing is more desirable when it is adaptive and still desirable when compared to note-taking. *Frontiers in psychology, 9*, 2596.
Herman, R. L. (2012). Letter from the Editor-in-Chief: the MOOCs are coming. *The Journal of Effective Teaching, 12*(2), 1-3.
Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational research review, 12*, 45-58.
Hibbing, A. N., & Rankin-Erickson, J. L. (2003). A picture is worth a thousand words: Using visual images to improve comprehension for middle school struggling readers. *The reading teacher, 56*(8), 758-770.
Hollands, F. M., & Tirthali, D. (2014). Resource requirements and costs of developing and delivering MOOCs. *International Review of Research in Open and Distributed Learning, 15*(5), 113-133.

Ibrahim, M., Antonenko, P. D., Greenwood, C. M., & Wheeler, D. (2012). Effects of segmenting, signalling, and weeding on learning from educational video. *Learning, Media and Technology, 37*(3), 220-235.

Iepsen, E. F., Bercht, M., & Reategui, E. (2013). *Detection and assistance to students who show frustration in learning of algorithms.* Paper presented at the 2013 IEEE Frontiers in Education Conference (FIE).

Iniesto, F., McAndrew, P., Minocha, S., & Coughlan, T. (2016). Accessibility of MOOCs: Understanding the Provider Perspective. *Journal of Interactive Media in Education, 2016*(1).

Jacob, W. J., Xiong, W., & Ye, H. (2015). Professional development programmes at world-class universities. *Palgrave Communications, 1*, 15002.

Jaffar, A. A. (2012). YouTube: An emerging tool in anatomy education. *Anatomical sciences education, 5*(3), 158-164.

Jansen, D., & Schuwer, R. (2015). Institutional MOOC strategies in Europe. *Status Report Based on a Mapping Survey Conducted in October-December 2014*.

Jeffries, M. (2009). Research in distance education. *MA Distributed Learning*.

Kaplan, A. M., & Haenlein, M. (2016). Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster. *Business Horizons, 59*(4), 441-450.

Keramida, M. (2015). What Is Wrong With MOOCs? Key Issues To Consider Before Launching Your First MOOC. *e_Learning Industry*.

Khalil, H., & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention-A Literature Review.

Kirkham, E. (2018). Study: Here's how much college credits actually cost. *StudentLoanHero. com, January, 24*.

Leito, I., Helm, I., & Jalukse, L. (2015). Using MOOCs for teaching analytical chemistry: experience at University of Tartu. *Analytical and bioanalytical chemistry, 407*(5), 1277-1281.

Li, X., Xu, H., Zhang, J., & Chang, H.-h. (2020). Deep Reinforcement Learning for Adaptive Learning Systems. *arXiv preprint arXiv:2004.08410*.

Liu, M., Kang, J., Cao, M., Lim, M., Ko, Y., Myers, R., & Schmitz Weiss, A. (2014). Understanding MOOCs as an emerging online learning tool: Perspectives from the students. *American Journal of Distance Education, 28*(3), 147-159.

Malcom-Piqueux, L., & Bensimon, E. M. (2017). Taking equity-minded action to close equity gaps. *Peer Review, 19*(2), 5.

Mandler, G. (1975). Mind and Emotion ‘. New York: John Wiley. *MandlerMind and Emotion1975*.

Marek, J., Stuchlíková, Ľ., Donoval, D., Chvála, A., Molnár, M., & Príbytný, P. (2015). Computer modelling and simulation of power mosfet properties in interactive e-learning course. *Distance Learning, Simulation and Communication 2015. Proceedings*, 92-97.

Marginson, S. (2016). Higher education and growing inequality.

Marques, J. F., & McCall, C. (2005). The application of interrater reliability as a solidification instrument in a phenomenological study. *The Qualitative Report, 10*(3), 439-462.

Marsh, E. J., & Butler, A. C. (2013). Memory in educational settings. *Oxford handbook of cognitive psychology*.

Mast, K. (2015). Multimedia in E-Learning.

Mayer, R. E. (2014). Incorporating motivation into multimedia learning. *Learning and instruction, 29*, 171-173.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of educational psychology, 100*(2), 380.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist, 38*(1), 43-52.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*: Sage publications.

Michigan, U. o. (2020). Ethnicity Reports. Retrieved from https://ro.umich.edu/reports/ethnicity

Miles, M. B., Huberman, A. M., & Saldaña, J. (1994). *Qualitative data analysis: An expanded sourcebook*: sage.

Mitchell, J. (2016). More Than 40% of Student Borrowers Aren't Making Payments. Retrieved from https://www.wsj.com/articles/more-than-40-of-student-borrowers-arent-making-payments-1459971348

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological science, 25*(6), 1159-1168.

Nanavati, A. A., & Bias, R. G. (2005). Optimal line length in reading-A literature review. *Visible Language, 39*(2), 120.

Newton, R. R., & Rudestam, K. E. (1999). Questions about multiple regression analysis. *Your statistical consultant: answers to your data analysis questions*, 247-275.

O'Brien, K., Forte, M., Mackey, T., & Jacobson, T. (2017). Metaliteracy as pedagogical framework for learner-centered design in three MOOC platforms: Connectivist, Coursera and Canvas. *Open Praxis, 9*(3), 267-286.

Onah, D. F., Sinclair, J., & Boyatt, R. (2014). Dropout rates of massive open online courses: behavioural patterns. *EDULEARN14 proceedings, 1*, 5825-5834.

Paas, F. G., & Van Merriënboer, J. J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of educational psychology, 86*(1), 122.

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., & Gowda, S. M. (2013). *Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes.* Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge.

Parr, C. (2012). Mooc completion rates 'below 7%'. Open online courses' cohort much less massive at finish line. *Times of higer education, 3*.

Peltier, J. W., Hay, A., & Drago, W. (2005). The reflective learning continuum: Reflecting on reflection. *Journal of marketing education, 27*(3), 250-263.

Peters, L. H., O'Connor, E. J., & Rudolf, C. J. (1980). The behavioral and affective consequences of performance-relevant situational variables. *Organizational Behavior and Human Performance, 25*(1), 79-96.

Pham, P., & Wang, J. (2017). *AttentiveLearner 2: a multimodal approach for improving MOOC learning on mobile devices.* Paper presented at the International Conference on Artificial Intelligence in Education.

Phobun, P., & Vicheanpanya, J. (2010). Adaptive intelligent tutoring systems for e-learning systems. *Procedia-Social and Behavioral Sciences, 2*(2), 4064-4069.

Pomales-Garcia, C., & Liu, Y. (2006). Web-based distance learning technology: The impacts of web module length and format. *The American journal of distance education, 20*(3), 163-179.

Radel, R., Pelletier, L., Baxter, D., Fournier, M., & Sarrazin, P. (2014). The paradoxical effect of controlling context on intrinsic motivation in another activity. *Learning and instruction, 29*, 95-102.

Reddy, S., Levine, S., & Dragan, A. (2017). *Accelerating human learning with deep reinforcement learning.* Paper presented at the NIPS workshop: teaching machines, robots, and humans.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science, 17*(3), 249-255.

Rogers, R. R. (2001). Reflection in higher education: A concept analysis. *Innovative higher education, 26*(1), 37-57.

Rosen, Y., Rushkin, I., Rubin, R., Munson, L., Ang, A., Weber, G., . . . Tingley, D. (2018). *The effects of adaptive learning in a massive open online course on learners' skill development.* Paper presented at the Proceedings of the Fifth Annual ACM Conference on Learning at Scale.

Rossiter, A., Nortcliffe, A., Griffin, A., & Middleton, A. (2009). Using student generated audio to enhance learning. *engineering education, 4*(2), 52-61.

Rout, S. K., Sahoo, M. P., & Das, M. J. (2020). Completion Rate of MOOCs by Pupil Teachers in DIETs in relation to their Self Directed Learning and Demographic Variables. *Studies in Indian Place Names, 40*(40), 1506-1516.

Sari, A. R., Bonk, C. J., & Zhu, M. (2020). MOOC instructor designs and challenges: what can be learned from existing MOOCs in Indonesia and Malaysia? *Asia Pacific Education Review, 21*(1), 143-166.

Schlosser, C. (1996). *Distance education: What the literature says works.* Paper presented at the Technology-Based Re-Engineering Engineering Education Proceedings of Frontiers in Education FIE'96 26th Annual Conference.

Schweppe, J., & Rummer, R. (2016). Integrating written text and graphics as a desirable difficulty in long-term multimedia learning. *Computers in Human Behavior, 60*, 131-137.

Shigeta, K., Koizumi, M., Sakai, H., Tsuji, Y., Inaba, R., & Hiraoka, N. (2017). A survey of the awareness, offering, and adoption of OERs and MOOCs in Japan. *Open Praxis, 9*(2), 195-206.

Singh, P. (2020). Challenges, Opportunities and Future Of MOOCs. *Studies in Indian Place Names, 40*(33), 97-100.

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*(4), 342-353.

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science, 10*(2), 176-199.

Stevens, V. (2013). What's with the MOOCs? *TESL-EJ, 16*(4).

Suzuki, Y., Nakata, T., & Dekeyser, R. (2019). The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice. *The Modern Language Journal, 103*(3), 713-720.

Swanwick, T. (2013). Understanding medical education. *Understanding medical education: evidence, theory and practice*, 1-6.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*(2), 257-285.
Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of educational psychology, 81*(4), 457.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction, 4*(4), 295-312.
Thomson, A., Bridgstock, R., & Willems, C. (2014). " Teachers Flipping Out" beyond the Online Lecture: Maximising the Educational Potential of Video. *Journal of Learning Design, 7*(3), 67-78.
Veletsianos, G. (2013). Learner experiences with MOOCs and open online learning. *Hybrid Pedagogy*.
Walia, K. (2020). BENEFITS AND CHALLENGES OF MOOCs: AN OVERVIEW. *Studies in Indian Place Names, 40*(33), 18-20.
Weber, G. (2012). Adaptive Learning Systems. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 113-115). Boston, MA: Springer US.
Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). An introduction to human factors engineering.
Williams, J. J., Kim, J., Rafferty, A., Maldonado, S., Gajos, K. Z., Lasecki, W. S., & Heffernan, N. (2016). *Axis: Generating explanations at scale with learnersourcing and machine learning.* Paper presented at the Proceedings of the Third (2016) ACM Conference on Learning@ Scale.
Wright, K. A., Lam, D. H., & Brown, R. G. (2009). Reduced approach motivation following nonreward: Extension of the BIS/BAS scales. *Personality and Individual Differences, 47*(7), 753-757.
Wu, W., Daskalakis, C., Kaashoek, N., Tzamos, C., & Weinberg, M. (2015). *Game theory based peer grading mechanisms for MOOCs.* Paper presented at the Proceedings of the Second (2015) ACM Conference on Learning@ Scale.
Xiao, X., Pham, P., & Wang, J. (2017). *Dynamics of affective states during mooc learning.* Paper presented at the International Conference on Artificial Intelligence in Education.
Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education, 140*, 103599.
Xing, B., & Marwala, T. (2017). Implications of the Fourth Industrial Age on Higher Education. *arXiv preprint arXiv:1703.09643*.
Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker Jr, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & management, 43*(1), 15-27.
Zhang, M., Zhu, J., Wang, Z., & Chen, Y. (2018). Providing personalized learning guidance in MOOCs by multi-source data analysis. *World Wide Web*, 1-31.
Zhang, Y. (2013). *Benefiting from MOOC.* Paper presented at the EdMedia+ Innovate Learning.
Zheng, B., Lin, C.-H., & Kwon, J. B. (2020). The impact of learner-, instructor-, and course-level factors on online learning. *Computers & Education, 150*, 103851.